

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification<sup>6</sup> :</b> <b>C07H 21/04</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 97/19098</b> <b>(43) International Publication Date:</b> 29 May 1997 (29.05.97)
<b>(21) International Application Number:</b> PCT/US96/18542 <b>(22) International Filing Date:</b> 15 November 1996 (15.11.96) <b>(30) Priority Data:</b> 08/561,469 17 November 1995 (17.11.95) US <b>(71) Applicant (for all designated States except US):</b> ASTRA AKTIEBOLAG [SE/SE]; S-151 85 Södertälje (SE). <b>(72) Inventor; and</b> <b>(75) Inventor/Applicant (for US only):</b> SMITH, Douglas, H. [US/US]; 2 Mayflower Lane, Gloucester, MA 01930 (US). <b>(74) Agents:</b> HANLEY, Elizabeth, A. et al.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the</i> <i>claims and to be republished in the event of the receipt of</i> <i>amendments.</i>
<b>(54) Title:</b> NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO HELICOBACTER PYLORI FOR DIAGNOSTICS AND THERAPEUTICS  <b>(57) Abstract</b>  Recombinant or substantially pure preparations of <i>H. pylori</i> polypeptides are described. The nucleic acids encoding the polypeptides also are described. The <i>H. pylori</i> polypeptides are useful in vaccine compositions.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

5 NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO *HELICOBACTER*  
*PYLORI* FOR DIAGNOSTICS AND THERAPEUTICS

Background of the Invention

*Helicobacter pylori* is a gram-negative, S-shaped, microaerophilic bacterium that  
10 was discovered and cultured from a human gastric biopsy specimen. (Warren, J.R. and B.  
Marshall, (1983) *Lancet* 1: 1273-1275; and Marshall et al., (1984) *Microbios Lett.* 25: 83-  
88). *H. pylori* has been strongly linked to chronic gastritis and duodenal ulcer disease.  
(Rathbone et. al., (1986) *Gut* 27: 635-641). Moreover, evidence is accumulating for an  
etiologic role of *H. pylori* in nonulcer dyspepsia, gastric ulcer disease, and gastric  
15 adenocarcinoma. (Blaser M. J., (1993) *Trends Microbiol.* 1: 255-260). Transmission of the  
bacteria occurs via the oral route, and the risk of infection increases with age. (Taylor,  
D.N. and M. J. Blaser, (1991) *Epidemiol. Rev* 13: 42-50). *H. pylori* colonizes the human  
gastric mucosa, establishing an infection that usually persists for decades. Infection by *H.*  
*pylori* is prevalent worldwide. Developed countries have infection rates over 50% of the  
20 adult population, while developing countries have infection rates reaching 90% of the  
adults over the age of 20. (Hopkins R. J. and J. G. Morris (1994) *Am. J. Med.* 97: 265-277).

The bacterial factors necessary for colonization of the gastric environment, and for  
virulence of this pathogen, are poorly understood. Examples of the putative virulence  
factors include the following: urease, an enzyme that may play a role in neutralizing gastric  
25 acid pH (Eaton et al., (1991) *Infect. Immunol.* 59: 2470-2475; Ferrero, R.L. and A. Lee  
(1991) *Microb. Ecol. Hlth. Dis.* 4: 121-134; Labigne et al., (1991) *J. Bacteriol.* 173: 1920-  
1931); the bacterial flagellar proteins responsible for motility across the mucous layer.  
(Hazell et al., (1986) *J. Inf. Dis.* 153: 658-663; Leying et al., (1992) *Mol. Microbiol.* 6:  
2863-2874; and Haas et al., (1993) *Mol. Microbiol.* 8: 753-760); Vac A, a bacterial toxin  
30 that induces the formation of intracellular vacuoles in epithelial cells (Schmitt, W. and R.  
Haas, (1994) *Molecular Microbiol.* 12(2): 307-319); and several gastric tissue-specific  
adhesins. (Boren et al., (1993) *Science* 262: 1892-1895; Evans et al., (1993) *J. Bacteriol.*  
175: 674-683; and Falk et al., (1993) *Proc. Natl. Acad. Sci. USA* 90: 2035-203).

Numerous therapeutic agents are currently available that eradicate *H. pylori*  
35 infections *in vitro*. (Huesca et. al., (1993) *Zbl. Bakt.* 280: 244-252; Hopkins, R. J. and J. G.  
Morris, supra). However, many of these treatments are suboptimally effective *in vivo*  
because of bacterial resistance, altered drug distribution, patient non-compliance or poor  
drug availability. (Hopkins, R. J. and J. G. Morris, supra). Treatment with antibiotics  
combined with bismuth are part of the standard regime used to treat *H. pylori* infection.  
40 (Malfertheiner, P. and J. E. Dominguez-Munoz (1993) *Clinical Therapeutics* 15 Supp. B:  
37-48). Recently, combinations of proton pump inhibitors and a single antibiotic have been  
shown to ameliorate duodenal ulcer disease. (Malfertheiner, P. and J. E. Dominguez-  
Munoz supra). However, methods employing antibiotic agents can have the problem of the

-2-

- 5 emergence of bacterial strains which are resistant to these agents. (Hopkins, R. J. and J. G. Morris, supra). These limitations demonstrate that new more effective methods are needed to combat *H. pylori* infections *in vivo*. In particular, the design of new vaccines that may prevent infection by this bacterium is highly desirable.

10

### Summary of the Invention

This invention relates to novel genes, e.g., genes encoding bacterial surface proteins, from the organism *Helicobacter pylori*, and other related genes, their products, and uses thereof. The nucleic acids and peptides of the present invention have utility for diagnostic and therapeutics for *H. pylori* and other *Helicobacter* species. They can also be used to detect the presence of *H. pylori* and other *Helicobacter* species in a sample; and for use in screening compounds for the ability to interfere with the *H. pylori* life cycle or to inhibit *H. pylori* infection. More specifically, this invention features compositions of nucleic acids corresponding to entire coding sequences of *H. pylori* surface proteins or parts thereof, nucleic acids capable of binding mRNA from *H. pylori* surface proteins to block protein translation, and methods for producing *H. pylori* surface proteins or parts thereof using peptide synthesis and recombinant DNA techniques. This invention also includes antibodies and nucleic acids sequences useful as probes to detect *H. pylori* infection. In addition, vaccine compositions and methods for the protection against infection by *H. pylori* are described.

25

### Brief Description of the Drawings

Figure 1 is a table which contains information from homology searches performed on the sequences of this invention using the BLAST algorithm contained in the Wisconsin Sequence Analysis Package.

30

Figure 2 is a bar graph that depicts the antibody titer in serum of mice following immunization with specific *H. pylori* antigens.

Figure 3 is a bar graph that depicts the antibody titer in mucous of mice following immunization with specific *H. pylori* antigens.

Figure 4 is a bar graph that depicts therapeutic immunization of *H. pylori* infected mice with specific antigens dissolved in HEPES buffer.

35

Figure 5 is a bar graph that depicts therapeutic immunization of *H. pylori* infected mice with specific antigens dissolved in buffer containing DOC.

### Detailed Description of the Invention

40

In one aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:115, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:1.



5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:116, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:2.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:117, such as a  
10 nucleic acid comprising a nucleotide sequence of SEQ ID NO:3.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:118, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:4.

In another aspect, the invention features a substantially pure nucleic acid encoding  
15 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:119, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:5.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:120, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:6.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:121, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:7.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:122, such as a  
25 nucleic acid comprising a nucleotide sequence of SEQ ID NO:8.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:123, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:9.

In another aspect, the invention features a substantially pure nucleic acid encoding  
30 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:124, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:10.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:125, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:11.

35 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:126, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:12.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:127, such as a  
40 nucleic acid comprising a nucleotide sequence of SEQ ID NO:13.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:128, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:14.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:129, such as a  
10 nucleic acid comprising a nucleotide sequence of SEQ ID NO:15.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:130, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:16.

In another aspect, the invention features a substantially pure nucleic acid encoding  
15 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:131, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:17.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:132, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:18.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:133, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:19.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:134, such as a  
25 nucleic acid comprising a nucleotide sequence of SEQ ID NO:20.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:135, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:21.

In another aspect, the invention features a substantially pure nucleic acid encoding  
30 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:136, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:22.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:137, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:23.

35 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:138, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:24.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:139, such as a  
40 nucleic acid comprising a nucleotide sequence of SEQ ID NO:25.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:140, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:26.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:141, such as a  
10 nucleic acid comprising a nucleotide sequence of SEQ ID NO:27.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:142, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:28.

In another aspect, the invention features a substantially pure nucleic acid encoding  
15 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:143, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:29.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:144, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:30.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:145, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:31.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:146, such as a  
25 nucleic acid comprising a nucleotide sequence of SEQ ID NO:32.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:147, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:33.

In another aspect, the invention features a substantially pure nucleic acid encoding  
30 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:148, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:34.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:149, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:35.

35 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:150, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:36.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:151, such as a  
40 nucleic acid comprising a nucleotide sequence of SEQ ID NO:37.

5           In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:152, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:38.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:153, such as a  
10   nucleic acid comprising a nucleotide sequence of SEQ ID NO:39.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:154, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:40.

          In another aspect, the invention features a substantially pure nucleic acid encoding  
15   an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:155, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:41.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:156, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:42.

20           In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:157, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:43.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:158, such as a  
25   nucleic acid comprising a nucleotide sequence of SEQ ID NO:44.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:159, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:45.

          In another aspect, the invention features a substantially pure nucleic acid encoding  
30   an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:160, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:46.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:161, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:47.

35           In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:162, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:48.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:163, such as a  
40   nucleic acid comprising a nucleotide sequence of SEQ ID NO:49.

-7-

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:164, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:50.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:165, such as a  
10 nucleic acid comprising a nucleotide sequence of SEQ ID NO:51.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:166, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:52.

In another aspect, the invention features a substantially pure nucleic acid encoding  
15 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:167, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:53.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:168, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:54.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:169, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:55.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:170, such as a  
25 nucleic acid comprising a nucleotide sequence of SEQ ID NO:56.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:171, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:57.

In another aspect, the invention features a substantially pure nucleic acid encoding  
30 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:172, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:58.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:173, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:59.

35 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:174, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:60.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:175, such as a  
40 nucleic acid comprising a nucleotide sequence of SEQ ID NO:61.

5           In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:176, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:62.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:177, such as a  
10       nucleic acid comprising a nucleotide sequence of SEQ ID NO:63.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:178, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:64.

          In another aspect, the invention features a substantially pure nucleic acid encoding  
15       an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:179, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:65.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:180, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:66.

20       In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:181, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:67.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:182, such as a  
25       nucleic acid comprising a nucleotide sequence of SEQ ID NO:68.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:183, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:69.

          In another aspect, the invention features a substantially pure nucleic acid encoding  
30       an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:184, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:70.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:185, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:71.

35       In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:186, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:72.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:187, such as a  
40       nucleic acid comprising a nucleotide sequence of SEQ ID NO:73.

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:188, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:74.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:189, such as a  
10 nucleic acid comprising a nucleotide sequence of SEQ ID NO:75.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:190, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:76.

In another aspect, the invention features a substantially pure nucleic acid encoding  
15 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:191, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:77.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:192, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:78.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:193, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:79.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:194, such as a  
25 nucleic acid comprising a nucleotide sequence of SEQ ID NO:80.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:195, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:81.

In another aspect, the invention features a substantially pure nucleic acid encoding  
30 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:196, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:82.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:197, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:83.

35 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:198, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:84.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:199, such as a  
40 nucleic acid comprising a nucleotide sequence of SEQ ID NO:85.

5           In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:200, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:86.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:201, such as a  
10       nucleic acid comprising a nucleotide sequence of SEQ ID NO:87.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:202, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:88.

          In another aspect, the invention features a substantially pure nucleic acid encoding  
15       an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:203, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:89.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:204, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:90.

20       In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:205, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:91.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:206, such as a  
25       nucleic acid comprising a nucleotide sequence of SEQ ID NO:92.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:207, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:93.

          In another aspect, the invention features a substantially pure nucleic acid encoding  
30       an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:208, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:94.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:209, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:95.

35       In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:210, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:96.

          In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:211, such as a  
40       nucleic acid comprising a nucleotide sequence of SEQ ID NO:97.



-11-

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:212, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:98.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:213, such as a  
10 nucleic acid comprising a nucleotide sequence of SEQ ID NO:99.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:214, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:100.

In another aspect, the invention features a substantially pure nucleic acid encoding  
15 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:215, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:101.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:216, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:102.

20 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:217, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:103.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:218, such as a  
25 nucleic acid comprising a nucleotide sequence of SEQ ID NO:104.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:219, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:105.

In another aspect, the invention features a substantially pure nucleic acid encoding  
30 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:220, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:106.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:221, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:107.

35 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:222, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:108.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:223, such as a  
40 nucleic acid comprising a nucleotide sequence of SEQ ID NO:109.

-12-

5 In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:224, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:110.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:225, such as a  
10 nucleic acid comprising a nucleotide sequence of SEQ ID NO:111.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:226, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:112.

In another aspect, the invention features a substantially pure nucleic acid encoding  
15 an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:227, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:113.

In another aspect, the invention features a substantially pure nucleic acid encoding an *H. pylori* polypeptide having an amino acid sequence of SEQ ID NO:228, such as a nucleic acid comprising a nucleotide sequence of SEQ ID NO:114.

20 In another aspect, the invention comprises nucleic acids capable of binding mRNA of *H. pylori*. Such nucleic acid is capable of acting as anti-sense nucleic acid to control the translation of mRNA of *H. pylori*. A further aspect features a nucleic acid which is capable of binding specifically to an *H. pylori* nucleic acid. Such nucleic acid has utility as probes and as capture reagents.

25 In another aspect, the invention features an expression system comprising an open reading frame corresponding to *H. pylori* nucleic acid. The nucleic acid further comprises a control sequence compatible with an intended host. The expression system is useful for making polypeptides corresponding to *H. pylori* nucleic acid.

In another aspect, the invention features a cell transformed with the expression  
30 system to make *H. pylori* polypeptides.

In another aspect, the invention features a method of generating antibodies against *H. pylori* proteins which are capable of binding specifically to *H. pylori* proteins. Such antibody has utility as reagents for immunoassays to evaluate the abundance and distribution of *H. pylori*-specific antigens.

35 In another aspect, the invention features a method of generating vaccines for immunizing an individual against *H. pylori*. The method includes: immunizing a subject with an *H. pylori* protein, e.g., a surface protein, or portion thereof, and a pharmaceutically acceptable carrier. Such vaccines have therapeutic and prophylactic utilities.

In another aspect, the invention provides a method for generating a vaccine  
40 comprising a modified immunogenic *H. pylori* protein, e.g., a surface protein, or portion thereof, and a pharmacologically acceptable carrier.

-13-

5 In another aspect, the invention features a method of evaluating a compound, e.g. a polypeptide, e.g., a fragment of a host cell polypeptide, for the ability to bind an *H. pylori* polypeptide. The method includes: contacting the candidate compound with an *H. pylori* polypeptides and determining if the compound binds or otherwise interacts with an *H. pylori* polypeptide. Compounds which bind *H. pylori* are candidates as activators or  
10 inhibitors of the bacterial life cycle. These assays can be performed *in vitro* or *in vivo*.

In another aspect, the invention features a method of evaluating a compound, e.g. a polypeptide, e.g., a fragment of a host cell polypeptide, for the ability to bind an *H. pylori* nucleic acid, e.g., DNA or RNA. The method includes: contacting the candidate  
15 compound with an *H. pylori* nucleic acid and determining if the compound binds or otherwise interacts with an *H. pylori* polypeptide. Compounds which bind *H. pylori* are candidates as activators or inhibitors of the bacterial life cycle. These assays can be performed *in vitro* or *in vivo*.

The invention features, *H. pylori* polypeptides, preferably a substantially pure preparation of an *H. pylori* polypeptide, or a recombinant *H. pylori* polypeptide. In  
20 preferred embodiments: the polypeptide has biological activity; the polypeptide has an amino acid sequence at least 60%, 80%, 90%, 95%, 98%, or 99% homologous to an amino acid sequence contained in SEQ ID NOs:115-228; the polypeptide has an amino acid sequence essentially the same as an amino acid sequence in SEQ ID NOs:115-228; the polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acids in length; the polypeptide  
25 includes at least 5, preferably at least 10, more preferably at least 20, more preferably at least 50, 100, or 150 contiguous amino acids contained in SEQ ID NOs:115-228.

In preferred embodiments: the *H. pylori* polypeptide is encoded by a nucleic acid in SEQ ID NOs:1-114, or by a nucleic acid having at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homology with a nucleic acid of SEQ ID NOs:1-114.

30 In a preferred embodiment, the subject *H. pylori* polypeptide differs in amino acid sequence at 1, 2, 3, 5, 10 or more residues from a sequence in SEQ ID NOs:115-228. The differences, however, are such that: the *H. pylori* polypeptide exhibits an *H. pylori* biological activity, e.g., the *H. pylori* polypeptide retains a biological activity of a naturally occurring *H. pylori* enzyme.

35 In preferred embodiments, the polypeptide includes all or a fragment of an amino acid sequence contained in SEQ ID NOs:115-228; fused, in reading frame, to additional amino acid residues, preferably to residues encoded by genomic DNA 5' to the genomic DNA which encodes a sequence contained in SEQ ID NOs:115-228.

In yet other preferred embodiments, the *H. pylori* polypeptide is a recombinant  
40 fusion protein having a first *H. pylori* polypeptide portion and a second polypeptide portion, e.g., a second polypeptide portion having an amino acid sequence unrelated to *H. pylori*. The second polypeptide portion can be, e.g., any of glutathione-S-transferase, a

-14-

- 5 DNA binding domain, or a polymerase activating domain. In preferred embodiment the fusion protein can be used in a two-hybrid assay.

Polypeptides of the invention include those which arise as a result of alternative transcription events, alternative RNA splicing events, and alternative translational and postranslational events.

- 10 The invention includes an immunogen which includes an *H. pylori* polypeptide in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for said *H. pylori* polypeptide, e.g., a humoral response, an antibody response, or a cellular response. In preferred embodiments, the immunogen comprises an antigenic determinant from a protein contained in SEQ ID NOs:115-228.

- 15 In another aspect, the invention provides a substantially pure nucleic acid having a nucleotide sequence which encodes an *H. pylori* polypeptide. In preferred embodiments: the encoded polypeptide has biological activity the encoded polypeptide has an amino acid sequence at least 60%, 80%, 90%, 95%, 98%, or 99% homologous to an amino acid sequence contained in SEQ ID NOs:115-228; the encoded polypeptide has an amino acid  
20 sequence essentially the same as an amino acid sequence in SEQ ID NOs:115-228; the encoded polypeptide is at least 5, 10, 20, 50, 100, or 150 amino acids in length; the encoded polypeptide comprises at least 5, preferably at least 10, more preferably at least 20, more preferably at least 50, 100, or 150 contiguous amino acids contained in SEQ ID NOs:115-228.

- 25 In preferred embodiments: the nucleic acid is that of SEQ ID NOs:1-114; the nucleic acid is at least 60%, 70%, 80%, 90%, 95%, 98%, or 99% homologous with a nucleic acid sequence contained in SEQ ID NOs:1-114.

- In a preferred embodiment, the encoded *H. pylori* polypeptide differs in amino acid sequence at 1, 2, 3, 5, 10 or more residues, from a sequence in SEQ ID NOs:115-228. The  
30 differences, however, are such that: the *H. pylori* encoded polypeptide exhibits a *H. pylori* biological activity, e.g., the encoded *H. pylori* enzyme retains a biological activity of a naturally occurring *H. pylori*.

- In preferred embodiments, the encoded polypeptide includes all or a fragment of an amino acid sequence contained in SEQ ID NOs:115-228; fused, in reading frame, to  
35 additional amino acid residues, preferably to residues encoded by genomic DNA 5' to the genomic DNA which encodes a sequence contained in SEQ ID NOs:115-228.

- In preferred embodiments, the subject *H. pylori* nucleic acid will include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter or transcriptional enhancer sequence, operably linked to the *H. pylori* gene sequence, e.g., to  
40 render the *H. pylori* gene sequence suitable for expression in a recombinant host cell.

In yet a further preferred embodiment, the nucleic acid which encodes an *H. pylori* polypeptide of the invention, hybridizes under stringent conditions to a nucleic acid probe

-15-

5 corresponding to at least 12 consecutive nucleotides contained in SEQ ID NOs:1-114; more preferably to at least 20 consecutive nucleotides contained in SEQ ID NOs:1-114; more preferably to at least 40 consecutive nucleotides contained in SEQ ID NOs:1-114.

In a preferred embodiment, the nucleic acid encodes a peptide which differs by at least one amino acid residue from the sequences shown in SEQ ID NOs:115-228.

10 In a preferred embodiment, the nucleic acid differs by at least one nucleotide from a nucleotide sequence shown in SEQ ID NOs:1-114 which encodes amino acids shown in SEQ ID NOs:115-228.

In another aspect, the invention includes: a vector including a nucleic acid which encodes an *H. pylori*-like polypeptide, e.g., an *H. pylori* polypeptide; a host cell transfected with the vector; and a method of producing a recombinant *H. pylori*-like polypeptide, e.g.,  
15 an *H. pylori* polypeptide; including culturing the cell, e.g., in a cell culture medium, and isolating the *H. pylori*-like polypeptide, e.g., an *H. pylori* polypeptide, e.g., from the cell or from the cell culture medium.

In another aspect, the invention features, a purified recombinant nucleic acid having  
20 at least 50%, 60%, 70%, 80%, 90%, 95%, 98%, or 99% homology with a sequence contained in SEQ ID NOs:1-114.

The invention also provides a probe or primer which includes a substantially purified oligonucleotide. The oligonucleotide includes a region of nucleotide sequence which hybridizes under stringent conditions to at least 10 consecutive nucleotides of sense  
25 or antisense sequence contained in SEQ ID NOs:1-114, or naturally occurring mutants thereof. In preferred embodiments, the probe or primer further includes a label group attached thereto. The label group can be, e.g., a radioisotope, a fluorescent compound, an enzyme, and/or an enzyme co-factor. Preferably the oligonucleotide is at least 10 and less than 20, 30, 50, 100, or 150 nucleotides in length.

30 The invention involves nucleic acids, e.g., RNA or DNA, encoding a polypeptide of the invention. This includes double stranded nucleic acids as well as coding and antisense single strands.

The *H. pylori* strain, from which genomic sequences have been sequenced, has been deposited in the American Type Culture Collection(ATCC # 55679) as strain HP-J99.

35 The nucleic acid sequences of this invention may be obtained directly from the DNA of the above referenced *H. pylori* strain by using the polymerase chain reaction (PCR). See "*PCR, A Practical Approach*" (McPherson, Quirke, and Taylor, eds., IRL Press, Oxford, UK, 1991) for details about the PCR. High fidelity PCR can be used to ensure a faithful DNA copy prior to expression. In addition, amplified products can be checked by  
40 conventional sequencing methods. Clones carrying the desired sequences described in this invention may be obtained by screening the libraries by means of the PCR or by hybridization of synthetic oligonucleotide probes to filter lifts of the library colonies or

-16-

5 plaques as known in the art (see, eg, Sambrook et al., *Molecular Cloning, A Laboratory Manual* 2nd edition, 1989, Cold Spring Harbor Press, NY).

Nucleic acids isolated or synthesized in accordance with features of the present invention are useful, by way of example, without limitation, as probes, primers, capture ligands, antisense genes and for developing expression systems for the synthesis of proteins and peptides corresponding to such sequences.

As probes, primers, capture ligands and antisense agents, the nucleic acid will normally comprise approximately twenty or more nucleotides for specificity as well as the ability to form stable hybridization products.

Putative functions have been determined for several of the *H. pylori* polypeptides of the invention, as shown in Figure 1.

Accordingly, uses of the claimed *H. pylori* polypeptides in these identified functions are also within the scope of the invention.

In addition, the present invention encompasses *H. pylori* polypeptides characterized as shown in Table 1 below, including: *H. pylori* outer membrane proteins, *H. pylori* periplasmic/secreted proteins, and other *H. pylori* surface proteins. Members of these groups were identified by BLAST homology searches. The *H. pylori* polypeptides identified in Table 1 are representative members of the groups identified above and are in no way limiting. Additional members of the groups can be identified within the *H. pylori* polypeptides disclosed herein by the methods known to those skilled in the art.

TABLE 1

SEQ ID NO	Blast identifier	Gene Symbol/Name	Description
<b>Outer Membrane Proteins</b>			
7116626 (SEQ ID NO:223)	P26093	e (P4)	e (P4) lipoprotein attached by lipid in <i>H. influenza</i>
29479681 (SEQ ID NO:179)	P13036	fecA	Receptor in Iron (III) dicitrate transport <i>E. coli</i>
36126938 (SEQ ID NO:199)	L12346	copB	Major out. memb. prot. in <i>M. catarrhalis</i>
<b>Periplasmic/Secreted Proteins</b>			
30100332 (SEQ ID NO:181)	P23847	dppA	Periplasmic dipeptide binding protein in <i>E. coli</i>
<b>Other Surface Proteins</b>			
4821082 (SEQ ID NO:212)	P08089	M protein	M protein of group A. <i>Streptococci</i>
978477 (SEQ ID NO:228)	L28919	FBP54	Surface Ag of grp A. <i>Streptococci</i> binds fibronectin

### Definitions

A purified preparation or a substantially pure preparation of a polypeptide, as used herein, means a polypeptide that has been separated from other proteins, lipids, and nucleic

-17-

5 acids with which it naturally occurs. Preferably, the polypeptide is also separated from substances, e.g., antibodies or gel matrix, e.g., polyacrylamide, which are used to purify it. Preferably, the polypeptide constitutes at least 10, 20, 50 70, 80 or 95% dry weight of the purified preparation. Preferably, the preparation contains: sufficient polypeptide to allow protein sequencing; at least 1, 10, or 100  $\mu$ g of the polypeptide; at least 1, 10, or 100 mg of  
10 the polypeptide.

A purified preparation of cells refers to, in the case of plant or animal cells, an in vitro preparation of cells and not an entire intact plant or animal. In the case of cultured cells or microbial cells, it consists of a preparation of at least 10% and more preferably 50% of the subject cells.

15 The metabolism of a substance, as used herein, means any aspect of the, expression, function, action, or regulation of the substance. The metabolism of a substance includes modifications, e.g., covalent or non covalent modifications of the substance. The metabolism of a substance includes modifications, e.g., covalent or non covalent modification, the substance induces in other substances. The metabolism of a substance  
20 also includes changes in the distribution of the substance. The metabolism of a substance includes changes the substance induces in the distribution of other substances.

A substantially pure nucleic acid, e.g., a substantially pure DNA, is a nucleic acid which is one or both of: not immediately contiguous with both of the coding sequences with which it is immediately contiguous (i.e., one at the 5' end and one at the 3' end) in the  
25 naturally-occurring genome of the organism from which the nucleic acid is derived; or which is substantially free of a nucleic acid sequence with which it occurs in the organism from which the nucleic acid is derived. The term includes, for example, a recombinant DNA which is incorporated into a vector, e.g., into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate  
30 molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other DNA sequences. Substantially pure DNA also includes a recombinant DNA which is part of a hybrid gene encoding additional *H. pylori* DNA sequence.

Homologous refers to the sequence similarity or sequence identity between two  
35 polypeptide molecules or between two nucleic acid molecules. When a position in both of the two compared sequences is occupied by the same base or amino acid monomer subunit, e.g., if a position in each of two DNA molecules is occupied by adenine, then the molecules are homologous at that position. The percent of homology between two sequences is a function of the number of matching or homologous positions shared by the two sequences  
40 divided by the number of positions compared x 100. For example, if 6 of 10, of the positions in two sequences are matched or homologous then the two sequences are 60% homologous. By way of example, the DNA sequences ATTGCC and TATGGC share 50%

-18-

5 homology. Generally, a comparison is made when two sequences are aligned to give maximum homology.

The terms peptides, proteins, and polypeptides are used interchangeably herein.

As used herein, the term "surface protein" refers to all surface accessible proteins, e.g. inner and outer membrane proteins, proteins adhering to the cell wall, and secreted  
10 proteins.

As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., one or more polypeptides), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be  
15 inserted, or is inserted, into the cell's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal expression of the selected nucleic acid, all operably linked to the  
20 selected nucleic acid, and may include an enhancer sequence.

As used herein, the term "transgenic cell" refers to a cell containing a transgene.

As used herein, a "transgenic animal" is any animal in which one or more, and preferably essentially all, of the cells of the animal includes a transgene. The transgene can be introduced into the cell, directly or indirectly by introduction into a precursor of the cell,  
25 by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. This molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

As used herein, the term "cell-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence operably  
30 linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

A polypeptide has *H. pylori* biological activity if it has one, two, three, and  
35 preferably more of the following properties: (1) if when expressed in the course of an *H. pylori* infection, it can promote, or mediate the attachment of *H. pylori* to a cell (2) it has an enzymatic activity characteristic of an *H. pylori* protein (3) or the gene which encodes it can rescue a lethal mutation in an *H. pylori* gene. A polypeptide has biological activity if it is an antagonist, agonist, or super-agonist of a polypeptide having one of the above-listed  
40 properties.

Misexpression, as used herein, refers to a non-wild type pattern of gene expression. It includes: expression at non-wild type levels, i.e., over or under expression; a pattern of



-19-

5 expression that differs from wild type in terms of the time or stage at which the gene is expressed, e.g., increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage; a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the  
10 splicing size, amino acid sequence, post-translational modification, or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene, e.g., a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus.

15 As used herein, "host cells" and other such terms denoting microorganisms or higher eukaryotic cell lines cultured as unicellular entities refers to cells which can become or have been used as recipients for a recombinant vector or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood by individuals skilled in the art that the progeny of a single parental cell may not necessarily be  
20 completely identical in genomic or total DNA complement to the original parent, due to accident or deliberate mutation.

As used herein, the term "control sequence" refers to a nucleic acid having a base sequence which is recognized by the host organism to effect the expression of encoded sequences to which they are ligated. The nature of such control sequences differs depending  
25 upon the host organism; in prokaryotes, such control sequences generally include a promoter, ribosomal binding site and terminators; in eukaryotes, generally such control sequences include promoters, terminators and in some instances, enhancers. The term control sequence is intended to include at a minimum, all components whose presence is necessary for expression, and may also include additional components whose presence is  
30 advantageous, for example, leader sequences.

As used herein, the term "operably linked" refers to sequences joined or ligated to function in their intended manner. For example, a control sequence is operably linked to coding sequence by ligation in such a way that expression of the coding sequence is achieved under conditions compatible with the control sequence and host cell.

35 An "open reading frame", also referred to herein as ORF, is a region of nucleic acid which encodes a peptide. This region may represent a portion of a coding sequence or a total sequence.

As used herein, a "coding sequence" is a nucleic acid sequence which is transcribed into messenger RNA and/or translated into a peptide when placed under the control of  
40 appropriate regulatory sequences. The boundaries of the coding sequence are determined by a translation start codon at the five prime terminus and a translation stop code at the three

-20-

5 prime terminus. A coding sequence can include but is not limited to messenger RNA, synthetic DNA, or recombinant nucleic acid sequences.

A "gene product" is a protein or structural RNA which is specifically encoded for by a gene.

10 As used herein, the term "probe" refers to a nucleic acid, peptide or other chemical entity which specifically binds to a molecule of interest. Probes are often associated with or capable of associating with a label. A label is a chemical moiety capable of detection. Typical labels comprise dyes, radioisotopes, luminescent and chemiluminescent moieties, fluorophores, enzymes, precipitating agents, amplification sequences, and the like. Similarly, a nucleic acid, peptide or other chemical entity which specifically binds to a  
15 molecule of interest and immobilizes such molecule is referred herein as a "capture ligand". Capture ligands are typically associated with or capable of associating with a support such as nitro-cellulose, glass, nylon membranes, beads, particles and the like. The specificity of hybridization is dependent on conditions such as the base pair composition of the nucleotides, and the temperature and salt concentration of the reaction. These conditions are  
20 readily discernable to one of ordinary skill in the art using routine experimentation.

The experimental manipulation of such conditions has been well described in the literature including such books as *Molecular Cloning; A Laboratory Manual*, Sambrook, J., Fritsch, E.F., Maniatis, T., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2nd ed. (1989).

25 The practice of the present invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See e.g., Sambrook, Fritsch, and Maniatis, *Molecular Cloning; Laboratory Manual* 2nd ed. (1989); *DNA Cloning*, Volumes I and II (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed. 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. 1984); the series, *Methods in Enzymology* (Academic Press, Inc.), particularly Vol. 154 and Vol. 155 (Wu and Grossman, eds.) and *PCR-A Practical Approach* (McPherson, Quirke, and Taylor, eds., 1991).

### 35 Probes

A nucleic acid isolated or synthesized in accordance with SEQ ID NOs:1-114 can be used as a probe to specifically detect *H. pylori*. With the sequence information set forth in the present application, sequences of twenty or more nucleotides are identified which provide the desired inclusivity and exclusivity with respect to *H. pylori*, and extraneous  
40 nucleic acid sequences likely to be encountered during hybridization conditions. More preferably, the sequence will comprise at least twenty to thirty nucleotides to convey

-21-

5 stability to the hybridization product formed between the probe and the intended target molecules.

Sequences larger than 1000 nucleotides in length are difficult to synthesize but can be generated by recombinant DNA techniques. Individuals skilled in the art will readily recognize that the nucleic acid sequences, for use as probes, can be provided with a label to  
10 facilitate detection of a hybridization product.

Nucleic acid isolated and synthesized in accordance with SEQ ID NOs:1-114 may also be useful as probes to detect homologous regions (especially homologous genes) of other *Helicobacter* species using relaxed stringency hybridization conditions, as will be obvious to anybody skilled in the art.

15

#### Capture Ligand

For use as a capture ligand, the nucleic acid selected in the manner described above with respect to probes, can be readily associated with a support. The manner in which nucleic acid is associated with supports is well known. Nucleic acid having twenty or more  
20 nucleotides in a sequence contained in SEQ ID NOs:1-114 have utility to separate *H. pylori* nucleic acid from the nucleic acid of each other and other organisms. Nucleic acid having twenty or more nucleotides in a sequence shown in SEQ ID NOs:1-114 may also have utility to separate other *Helicobacter* species from each other and from other organisms. Preferably, the sequence will comprise at least twenty nucleotides to convey stability to the  
25 hybridization product formed between the probe and the intended target molecules. Sequences larger than 1000 nucleotides in length are difficult to synthesize but can be generated by recombinant DNA techniques.

#### Primers

30 Nucleic acid isolated or synthesized in accordance with the sequences described herein have utility as primers for the amplification of *H. pylori* nucleic acid. These nucleic acids may also have utility as primers for the amplification of nucleic acid sequences in other *Helicobacter* species. With respect to polymerase chain reaction (PCR) techniques, nucleic acid sequences of  $\geq 10$ -15 nucleotides contained in SEQ ID NOs:1-114 have utility  
35 in conjunction with suitable enzymes and reagents to create copies of *H. pylori* nucleic acid. More preferably, the sequence will comprise twenty or more nucleotides to convey stability to the hybridization product formed between the primer and the intended target molecules. Binding conditions of primers greater than 100 nucleotides are more difficult to control to obtain specificity. High fidelity PCR can be used to ensure a faithful DNA copy  
40 prior to expression. In addition, amplified products can be checked by conventional sequencing methods.

-22-

5           The copies can be used in diagnostic assays to detect specific sequences, including genes from *H. pylori* and/or other *Helicobacter* species. The copies can also be incorporated into cloning and expression vectors to generate polypeptides corresponding to the nucleic acid synthesized by PCR, as will be described in greater detail below.

10    Antisense

          Nucleic acid or nucleic acid-hybridizing derivatives isolated or synthesized in accordance with the sequences described herein have utility as antisense agents to prevent the expression of *H. pylori* genes. These sequences may also have utility as antisense agents to prevent expression of genes of other *Helicobacter* species.

15           Nucleic acid or derivatives corresponding to *H. pylori* nucleic acid sequences is loaded into a suitable carrier such as a liposome or bacteriophage for introduction into bacterial cells. For example, a nucleic acid having twenty or more nucleotides is capable of binding to bacteria nucleic acid or bacteria messenger RNA. Preferably, the antisense nucleic acid is comprised of 20 or more nucleotides to provide necessary stability of a  
20    hybridization product of non-naturally occurring nucleic acid and bacterial nucleic acid and/or bacterial messenger RNA. Nucleic acid having a sequence greater than 1000 nucleotides in length is difficult to synthesize but can be generated by recombinant DNA techniques. Methods for loading antisense nucleic acid in liposomes is known in the art as exemplified by U.S. Patent 4,241,046 issued December 23, 1980 to Papahadjopoulos et al.

25

Expressing *H. pylori* Genes

          The function of a specific gene or operon can be ascertained by expression in a bacterial strain under conditions where the activity of the gene product(s) specified by the gene or operon in question can be specifically measured. Alternatively, a gene product may  
30    be produced in large quantities in an expressing strain for use as an antigen, an industrial reagent, for structural studies, etc. This expression could be accomplished in a mutant strain which lacks the activity of the gene to be tested, or in a strain that does not produce the same gene product(s). This includes, but is not limited to other *Helicobacter* strains, and other bacterial strains such as *E. coli*, *Norcardia*, *Corynebacterium*, and *Streptomyces*  
35    species. In some cases the expression host will utilize the natural *Helicobacter* promoter whereas in others, it will be necessary to drive the gene with a promoter sequence derived from the expressing organism (e.g., an *E. coli* beta-galactosidase promoter for expression in *E. coli*).

40           To express a gene product using the natural *H. pylori* promoter, a procedure such as the following is used. A restriction fragment containing the gene of interest, together with its associated natural promoter element and regulatory sequences (identified using the DNA sequence data) is cloned into an appropriate recombinant plasmid containing the following

-23-

5 components: an origin of replication that functions in the host organism, and an appropriate selectable marker. This can be accomplished by a number of procedures known to those skilled in the art. It is most preferably done by cutting the plasmid and the fragment to be cloned with the same restriction enzyme to produce compatible ends that can be ligated to join the two pieces together. The recombinant plasmid is introduced into the host organism  
10 by electroporation and cells containing the recombinant plasmid are identified by selection for the marker on the plasmid. Expression of the desired gene product is detected using an assay specific for that gene product.

In the case of a gene that requires a different promoter, the body of the gene (coding sequence) is specifically excised and cloned into an appropriate expression plasmid. This  
15 subcloning can be done by several methods, but is most easily accomplished by PCR amplification of a specific fragment and ligation into an expression plasmid after treating the PCR product with a restriction enzyme or exonuclease to create suitable ends for cloning.

## 20 Expressed Genes in Therapeutics

Nucleic acid isolated or synthesized in accordance with the sequences described herein have utility to generate proteins and peptides. The nucleic acid exemplified in SEQ ID NOs:1-114 or fragments of said nucleic acid sequences encoding immunogenic portions of *H. pylori* proteins (SEQ ID NO:115-228) can be cloned into suitable vectors or used to  
25 isolate nucleic acid. The isolated nucleic acid is combined with suitable DNA linkers and cloned into a suitable vector.

The host cell may be any procaryotic or eucaryotic cell. For example, an *H. pylori* peptide may be expressed in bacterial cells such as *E. coli*, insect cells (baculovirus), yeast, or mammalian cells such as Chinese hamster ovary cell (CHO). Other suitable host cells  
30 are known to those skilled in the art.

Expression in eucaryotic cells such as mammalian, yeast, or insect cells can lead to partial or complete glycosylation and/or formation of relevant inter- or intra-chain disulfide bonds of a recombinant peptide product. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, et al., (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Baculovirus vectors available for expression of proteins in cultured insect cells (SF 9 cells) include the pAc series (Smith et al., (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow, V.A., and Summers, M.D., (1989) *Virology* 170:31-39). Generally, COS cells (Gluzman, Y., (1981) *Cell* 23:175-182) are used in conjunction with such vectors as pCDM 8 (Aruffo, A. and Seed, B., (1987) *Proc. Natl. Acad. Sci. USA* 84:8573-8577) for transient  
40 amplification/expression in mammalian cells, while CHO (dhfr<sup>-</sup> Chinese Hamster Ovary)

-24-

5 cells are used with vectors such as pMT2PC (Kaufman et al. (1987), *EMBO J.* 6:187-195) for stable amplification/expression in mammalian cells. Vector DNA can be introduced into mammalian cells via conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, or electroporation. Suitable methods for transforming host cells can be found in Sambrook et al. (Molecular  
10 Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory press (1989)), and other laboratory textbooks.

Expression in procaryotes is most often carried out in *E. coli* with either fusion or non-fusion inducible expression vectors. Fusion vectors usually add a number of NH<sub>2</sub> terminal amino acids to the expressed target gene. These NH<sub>2</sub> terminal amino acids often  
15 are referred to as a reporter group. Such reporter groups usually serve two purposes: 1) to increase the solubility of the target recombinant protein; and 2) to aid in the purification of the target recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the reporter group and the target recombinant protein to enable separation of the target recombinant  
20 protein from the reporter group subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Amrad Corp., Melbourne, Australia), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase, maltose E binding protein, or protein  
25 A, respectively, to the target recombinant protein. A preferred reporter group is poly(His), which may be fused to the amino or carboxy terminus of the protein and which renders the recombinant fusion protein easily purifiable by metal chelate chromatography.

Inducible non-fusion expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET11d (Studier et al., Gene Expression Technology: Methods in  
30 Enzymology 185, Academic Press, San Diego, California (1990) 60-89). While target gene expression relies on host RNA polymerase transcription from the hybrid trp-lac fusion promoter in pTrc, expression of target genes inserted into pET11d relies on transcription from the T7 gn10-lac 0 fusion promoter mediated by coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3)  
35 from a resident  $\lambda$  prophage harboring a T7 gn1 under the transcriptional control of the lacUV 5 promoter.

For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding an *H. pylori* peptide can be cultured under appropriate conditions to allow expression of the peptide to occur. The peptide may be secreted and  
40 isolated from a mixture of cells and medium containing the peptide. Alternatively, the peptide may be retained cytoplasmically and the cells harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for

-25-

5 cell culture are well known in the art. Peptides of the invention can be isolated from cell culture medium, host cells, or both using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such peptides. Additionally, in many situations, peptides can be produced by chemical cleavage of a  
10 native protein (e.g., tryptic digestion) and the cleavage products can then be purified by standard techniques.

In the case of membrane bound proteins, these can be isolated from a host cell by contacting a membrane-associated protein fraction with a detergent forming a solubilized complex, where the membrane-associated protein is no longer entirely embedded in the  
15 membrane fraction and is solubilized at least to an extent which allows it to be chromatographically isolated from the membrane fraction. Several different criteria are used for choosing a detergent suitable for solubilizing these complex. For example, one property considered is the ability of the detergent to solubilize the *H. pylori* protein within the membrane fraction at minimal denaturation of the membrane-associated protein  
20 allowing for the activity or functionality of the membrane-associated protein to return upon reconstitution of the protein. Another property considered when selecting the detergent is the critical micells concentration (CMC) of the detergent in that the detergent of choice preferably has a high CMC value allowing for ease of removal after reconstitution. A third property considered when selecting a detergent is the hydrophobicity of the detergent.  
25 Typically, membrane-associated proteins are very hydrophobic and therefore detergents which are also hydrophobic, e.g. the triton series, would be useful for solubilizing the hydrophobic proteins. Another property important to a detergent can be the capability of the detergent to remove the *H. pylori* protein with minimal protein-protein interaction facilitating further purification. A fifth property of the detergent which should be  
30 considered is the charge of the detergent. For example, if it is desired to use ion exchange resins in the purification process then preferably detergent should be an uncharged detergent. Chromatographic techniques which can be used in the final purification step are known in the art and include hydrophobic interaction, lectin affinity, ion exchange, dye affinity and immunoaffinity.

35 One strategy to maximize recombinant *H. pylori* peptide expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 119-128). Another strategy would be to alter the nucleic acid encoding *H. pylori* peptide to be inserted into an  
40 expression vector so that the individual codons for each amino acid would be those preferentially utilized in highly expressed *E. coli* proteins (Wada et al., (1992) *Nuc. Acids*

-26-

- 5     *Res.* 20:2111-2118). Such alteration of nucleic acids of the invention can be carried out by standard DNA synthesis techniques.

          The nucleic acids of the invention can also be chemically synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides are known, including solid-phase synthesis which, like peptide synthesis, has been fully  
10     automated in commercially available DNA synthesizers (See, e.g., Itakura et al. U.S. Patent No. 4,598,049; Caruthers et al. U.S. Patent No. 4,458,066; and Itakura U.S. Patent Nos. 4,401,796 and 4,373,071, incorporated by reference herein).

#### Drug Screening Assays

- 15           By making available purified and recombinant *H. pylori* polypeptides, the present invention provides assays which can be used to screen for drugs which are either agonists or antagonists of the normal cellular function, in this case, of the subject *H. pylori* polypeptides, or of their role in intracellular signaling. Such inhibitors or potentiators may be useful as new therapeutic agents to combat *H. pylori* infections in humans. A variety of  
20     assay formats will suffice and, in light of the present inventions, will be comprehended by the skilled artisan.

          In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays which are performed in cell-free  
25     systems, such as may be derived with purified or semi-purified proteins, are often preferred as "primary" screens in that they can be generated to permit rapid development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the *in vitro* system, the assay instead being focused  
30     primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with other proteins or change in enzymatic properties of the molecular target. Accordingly, in an exemplary screening assay of the present invention, the compound of interest is contacted with an isolated and purified *H. pylori* polypeptide.

          Screening assays may be constructed *in vitro* with a purified *H. pylori* enzyme such  
35     that the action of the enzyme produces an easily detectable reaction product. The efficacy of the compound can be assessed by generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. Suitable products include those with distinctive absorption, fluorescence, or chemi-luminescence properties, for example,  
40     because detection may be easily automated. A variety of synthetic or naturally occurring compounds may be tested in the assay to identify those which inhibit or potentiate the activity of the *H. pylori* enzyme. Some of these active compounds may directly, or with



-27-

- 5 chemical alterations to promote membrane permeability or solubility, also inhibit or potentiate the same enzymatic activity in whole, live *H. pylori* cells.

### Antibodies

- The invention also includes antibodies specifically reactive with the subject *H. pylori*-like polypeptide. Anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, *Antibodies: A Laboratory Manual* ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal such as a mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide. Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic portion of the subject *H. pylori* polypeptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of the *H. pylori* polypeptides of the invention, e.g. antigenic determinants of a polypeptide shown in SEQ ID NOs:115-228 or a closely related human or non-human mammalian homolog (e.g. 90% percent homologous, more preferably at least 95 percent homologous). In yet a further preferred embodiment of the present invention, the anti-*H. pylori* antibodies do not substantially cross react (i.e. react specifically) with a protein which is: e.g., less than 80% percent homologous to a sequence shown in SEQ ID NOs:115-228. By "not substantially cross react", it is meant that the antibody has a binding affinity for a non-homologous protein which is less than 10 percent, more preferably less than 5 percent, and even more preferably less than 1 percent, of the binding affinity for a protein contained in SEQ ID NOs:115-228. In a most preferred embodiment, there is no crossreactivity between bacterial and mammalian antigens.

- The term antibody as used herein is intended to include fragments thereof which are also specifically reactive with *H. pylori* polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab')<sub>2</sub> fragments can be generated by treating antibody with pepsin. The resulting F(ab')<sub>2</sub> fragment can be treated to reduce disulfide bridges to produce Fab' fragments. The antibody of the present invention is further intended to include bispecific and chimeric molecules having an anti-*H. pylori* portion.

- Both monoclonal and polyclonal antibodies (Ab) directed against *H. pylori* polypeptides or *H. pylori* polypeptide variants, and antibody fragments such as Fab' and F(ab')<sub>2</sub>, can be used to block the action of *H. pylori* polypeptide and allow the study of the role of a particular *H. pylori* polypeptide of the present invention in aberrant or unwanted

-28-

5 intracellular signaling, as well as the normal cellular function of the *H. pylori* and by microinjection of anti-*H. pylori* polypeptide antibodies of the present invention.

Antibodies which specifically bind *H. pylori* epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of *H. pylori* antigens. Anti-*H. pylori* polypeptide antibodies can be  
10 used diagnostically in immuno-precipitation and immuno-blotting to detect and evaluate *H. pylori* levels in tissue or bodily fluid as part of a clinical testing procedure. Likewise, the ability to monitor *H. pylori* polypeptide levels in an individual can allow determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. The level of an *H. pylori* polypeptide can be measured in cells found in bodily fluid, such  
15 as in urine samples or can be measured in tissue, such as produced by gastric biopsy. Diagnostic assays using anti-*H. pylori* antibodies can include, for example, immunoassays designed to aid in early diagnosis of *H. pylori* infections. The present invention can also be used as a method of detecting antibodies contained in samples from individuals infected by this bacterium using specific *H. pylori* antigens.

20 Another application of anti-*H. pylori* polypeptide antibodies of the present invention is in the immunological screening of cDNA libraries constructed in expression vectors such as  $\lambda$ gt11,  $\lambda$ gt18-23,  $\lambda$ ZAP, and  $\lambda$ ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion  
25 proteins. For instance,  $\lambda$ gt11 will produce fusion proteins whose amino termini consist of  $\beta$ -galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a subject *H. pylori* polypeptide can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with anti-*H. pylori* polypeptide antibodies. Phage, scored by this assay, can then be isolated from the infected plate. Thus, the presence of *H. pylori* gene homologs can be  
30 detected and cloned from other species, and alternate isoforms (including splicing variants) can be detected and cloned.

### Vaccines

The present invention also includes vaccine compositions for protection against  
35 infection by *H. pylori* or for treatment of *H. pylori* infection, a gram-negative spiral microaerophilic bacterium. In one embodiment, the vaccine compositions contain immunogenic surface proteins from *H. pylori*, or portion thereof, and a pharmaceutically acceptable carrier. Nucleic acids within the scope of the invention are exemplified by the nucleic acids shown in SEQ ID NOs:1-114 and which encode *H. pylori* surface proteins  
40 shown in SEQ ID NOs:115-228. However, any nucleic acid encoding an immunogenic *H. pylori* protein, or portion thereof, which is capable of expression in a cell, can be used in the present invention. These vaccines can have therapeutic and prophylactic utilities.

-29-

5 Another aspect of the present invention provides vaccine compositions for protection against infection by *H. pylori* or for treatment of *H. pylori* infection, which contain a modified immunogenic *H. pylori* protein or portion thereof, and a pharmaceutically acceptable carrier. It is possible to modify the structure of a *H. pylori* protein or peptide for such purposes as increasing solubility, enhancing stability (e.g., shelf  
10 life *ex vivo* and resistance to proteolytic degradation *in vivo*). A modified *H. pylori* protein or peptide can be produced in which the amino acid sequence has been altered, such as by amino acid substitution, deletion, or addition.

Another example of modification of an *H. pylori* peptide is substitution of cysteine residues preferably with alanine, serine, threonine, leucine or glutamic acid residues to  
15 minimize dimerization via disulfide linkages. In addition, amino acid side chains of fragments of the protein of the invention can be chemically modified. Another modification is cyclization of the peptide.

In order to enhance stability and/or reactivity, an *H. pylori* protein or peptide can be modified to incorporate one or more polymorphisms in the amino acid sequence of the  
20 protein resulting from any natural allelic variation. Additionally, D-amino acids, non-natural amino acids, or non-amino acid analogs can be substituted or added to produce a modified protein within the scope of this invention. Furthermore, an *H. pylori* protein can be modified using polyethylene glycol (PEG) according to the method of A. Schon and co-workers (Wie et al., *supra*) to produce a protein conjugated with PEG. In addition, PEG  
25 can be added during chemical synthesis of the protein. Other modifications of *H. pylori* proteins include reduction/alkylation (Tarr, *Methods of Protein Microcharacterization*, J. E. Silver ed., Humana Press, Clifton NJ 155-194 (1986)); acylation (Tarr, *supra*); chemical coupling to an appropriate carrier (Mishell and Shiigi, eds. *Selected Methods in Cellular Immunology*, WH Freeman, San Francisco, CA (1980), U.S. Patent 4,939,239; or mild  
30 formalin treatment (Marsh, (1971) *Int. Arch. of Allergy and Appl. Immunol.*, 41: 199 - 215).

To facilitate purification and potentially increase solubility of an *H. pylori* protein or peptide, it is possible to add an amino acid fusion moiety to the peptide backbone. For example, hexa-histidine can be added to the protein for purification by immobilized metal ion affinity chromatography (Hochuli, E. et al., (1988) *Bio/Technology*, 6: 1321 - 1325). In  
35 addition, to facilitate isolation of peptides free of irrelevant sequences, specific endoprotease cleavage sites can be introduced between the sequences of the fusion moiety and the peptide.

To potentially aid proper antigen processing of T cell epitopes within an *H. pylori* protein of the invention, canonical protease sensitive sites can be engineered between  
40 regions, each comprising at least one T cell epitope via recombinant or synthetic methods. For example, charged amino acid pairs, such as KK or RR, can be introduced between regions within a protein or fragment during recombinant construction thereof. The

-30-

5 resulting peptide can be rendered sensitive to cleavage by cathepsin and/or other trypsin-like enzymes which would generate portions of the protein containing one or more T cell epitopes. In addition, such charged amino acid residues can result in an increase in the solubility of the peptide.

10 Site-directed mutagenesis of a nucleic acid encoding an *H. pylori* protein can be used to modify the structure of the peptide by methods known in the art. Such methods may, among others, include polymerase chain reaction (PCR) with oligonucleotide primers bearing one or more mutations (Ho et al., (1989) *Gene*, 77: 51 - 59) or total synthesis of mutated genes (Hostomsky, Z. et al., (1989) *Biochem. Biophys. Res. Comm.*, 161: 1056 - 1063). To enhance recombinant protein expression, the aforementioned methods can be  
15 applied to change the codons present in the cDNA sequence of the invention to those preferentially utilized by the host cell in which the recombinant protein is being expressed (Wada et al., supra). An extensive discussion of mutagenesis protocols is provided in the "Production of Fragments and Analogs" section herein.

20 Another aspect of the invention provides a vaccine composition for protection against infection by *H. pylori* which contains an immunogenic fragment of an *H. pylori* protein or portion thereof, and a pharmaceutically acceptable carrier. Preferred fragments include peptides of at least about 10 amino acid residues in length, preferably about 10-20 amino acid residues in length, and more preferably about 12-16 amino acid residues in length.

25 Immunogenic peptides of the invention can be obtained, for example, by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid encoding the full-length *H. pylori* protein. In addition, fragments can be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, an *H. pylori* protein may be arbitrarily divided  
30 into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced (recombinantly or by chemical synthesis) and tested to identify those peptides having the ability to induce a T cell response, such as stimulation (proliferation, cytokine secretion). An extensive discussion of peptide analogs and fragments is provided in the "Production of  
35 Fragments and Analogs" section herein.

40 In one embodiment, immunogenic *H. pylori* fragments can be identified by the ability of the peptide to stimulate T cells. Peptides which stimulate T cells, as determined by, for example, T cell proliferation or cytokine secretion are defined herein as comprising at least one T cell epitope. T cell epitopes are believed to be involved in initiation and perpetuation of the immune response to the protein allergen which is responsible for the clinical symptoms of allergy. These T cell epitopes are thought to trigger early events at the level of the T helper cell by binding to an appropriate HLA molecule on the surface of

-31-

5 an antigen presenting cell, thereby stimulating the T cell subpopulation with the relevant T cell receptor for the epitope. These events lead to T cell proliferation, lymphokine secretion, local inflammatory reactions, recruitment of additional immune cells to the site of antigen/T cell interaction, and activation of the B cell cascade, leading to the production of antibodies. One isotype of these antibodies, IgE, is fundamentally important to the  
10 development of allergic symptoms and its production is influenced early in the cascade of events at the level of the T helper cell, by the nature of the lymphokines secreted. A T cell epitope is the basic element, or smallest unit of recognition by a T cell receptor, where the epitope comprises amino acids essential to receptor recognition. Amino acid sequences which mimic those of the T cell epitopes and which modify the allergic response to protein  
15 allergens are within the scope of this invention.

In another embodiment, immunogenic components of the invention are identified through genomic vaccination. The basic protocol is based on the idea that expression libraries consisting of all or parts of a pathogen genome, e.g., an *H. pylori* genome, can confer protection when used to genetically immunize a host. This expression library  
20 immunization (ELI) is analogous to expression cloning and involves reducing a genomic expression library of a pathogen, e.g., *H. pylori*, into plasmids that can act as genetic vaccines. The plasmids can also be designed to encode genetic adjuvants which can dramatically stimulate the humoral response. These genetic adjuvants can be introduced at remote sites and act as well extracellularly as intracellularly.

25 This is a new approach to vaccine production that has many of the advantages of live/attenuated pathogens but no risk of infection. An expression library of pathogen DNA is used to immunize a host thereby producing the effects of antigen presentation of a live vaccine without the risk. For example, in the present invention, random fragments from the *H. pylori* genome or from cosmid or plasmid clones, as well as PCR products from genes  
30 identified by genomic sequencing, can be used to immunize a host. The feasibility of this approach has been demonstrated with *Mycoplasma pulmonis* (Barry et al., *Nature* 377:632-635, 1995), where even partial expression libraries of *Mycoplasma pulmonis*, a natural pathogen in rodents, provided protection against challenge from the pathogen.

ELI is a technique that allows for production of a non-infectious multipartite  
35 vaccine, even when little is known about pathogen's biology, because ELI uses the immune system to screen candidate genes. Once isolated, these genes can be used as genetic vaccines or for development of recombinant protein vaccines. Thus, ELI allows for production of vaccines in a systematic, largely mechanized fashion.

Screening peptides for those which are immunogenic can be accomplished using  
40 one or more of several different assays. For example, *in vitro*, peptide T cell stimulatory activity is assayed by contacting a peptide known or suspected of being immunogenic with an antigen presenting cell which presents appropriate MHC molecules in a T cell culture.

-32-

5 Presentation of an immunogenic *H. pylori* peptide in association with appropriate MHC molecules, to T cells, in conjunction with the necessary costimulation, has the effect of transmitting a signal to the T cell that induces the production of increased levels of cytokines, particularly of interleukin-2 and interleukin-4. The culture supernatant can be obtained and assayed for interleukin-2 or other known cytokines. For example, any one of  
10 several conventional assays for interleukin-2 can be employed, such as the assay described in *Proc. Natl. Acad. Sci USA*, 86: 1333 (1989) the pertinent portions of which are incorporated herein by reference. A kit for an assay for the production of interferon is also available from Genzyme Corporation (Cambridge, MA).

Alternatively, a common assay for T cell proliferation entails measuring tritiated  
15 thymidine incorporation. The proliferation of T cells can be measured *in vitro* by determining the amount of <sup>3</sup>H-labeled thymidine incorporated into the replicating DNA of cultured cells. Therefore, the rate of DNA synthesis and, in turn, the rate of cell division can be quantified.

To produce modified proteins or immunogenic fragments by recombinant DNA  
20 techniques, an expression vector containing a nucleic acid encoding all or a portion of a *H. pylori* protein, operably linked to at least one regulatory sequence can be used. Operably linked is intended to mean that the nucleotide sequence is linked to a regulatory sequence in a manner which allows expression of the nucleotide sequence. Regulatory sequences are art-recognized and include promoters, enhancers and other expression control elements.  
25 Such regulatory sequences are described in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). It should be understood that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed and/or the type of protein desired to be expressed. In one embodiment of the present invention, the expression vector includes nucleic acid,  
30 preferably a DNA, encoding a modified *H. pylori* protein or immunogenic fragment having all or a portion of the amino acid sequence. Such expression vectors can be used to transfect cells to thereby produce proteins or peptides, including fusion proteins or peptides encoded by nucleic acids as described herein.

Host cells suitable for transfection and recombinant production of *H. pylori* proteins  
35 of the invention include any procaryotic or eucaryotic cell. For example, an *H. pylori* protein or peptide may be expressed in bacterial cells such as *E. coli*, insect cells (baculovirus), yeast, or mammalian cells such as Chinese hamster ovary cells (CHO). Other suitable host cells can be found in Goeddel, (1990) *supra* or known to those skilled in the art.

40 *H. pylori* proteins and fragments of the invention can also be chemically synthesized, using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. The nucleic acids of the invention can also be chemically

-33-

5 synthesized using standard techniques. Various methods of chemically synthesizing polydeoxynucleotides are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (see e.g., Itakura et al. U.S. Patent No. 4,598,049; Caruthers et al., U.S. Patent No. 4,458,066; and Itakura, U.S. Patent Nos. 4,401,796 and 4,373,071, incorporated by reference herein).

10 Vaccine compositions of the present invention containing DNA encoding immunogenic protein from *H. pylori*, or containing modified protein or fragments, contain both the DNA or protein and a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier that does not cause an allergic reaction or other untoward effect in patients to whom it is administered. Suitable  
15 pharmaceutically acceptable carriers include, for example, one or more of water, saline, phosphate buffered saline, dextrose, glycerol, ethanol and the like, as well as combinations thereof. Pharmaceutically acceptable carriers may further comprise minor amounts of auxiliary substances such as wetting or emulsifying agents, preservatives or buffers, which enhance the shelf life or effectiveness of the antibody. For vaccines of the invention  
20 containing modified *H. pylori* protein or immunogenic protein fragments, the protein or peptide must be coadministered with a suitable adjuvant.

It will be apparent to those of skill in the art that the therapeutically effective amount of DNA or protein of this invention will depend, *inter alia*, upon the administration schedule, the unit dose of antibody administered, whether the protein or DNA is  
25 administered in combination with other therapeutic agents, the immune status and health of the patient, and the therapeutic activity of the particular protein or DNA.

Vaccine compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Methods for intramuscular immunization are described by Wolff et al. (1990) *Science* 247: 1465-1468 and by Sedegah  
30 et al. (1994) *Immunology* 91: 9866-9870. Other modes of administration include oral and pulmonary formulations, suppositories, and transdermal applications. Oral immunization is preferred over parenteral methods for inducing protection against infection by *H. pylori*. Czinn et. al. (1993) *Vaccine* 11: 637-642. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch,  
35 magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like.

The vaccine compositions of the invention can include an adjuvant, including, but not limited to aluminum hydroxide; N-acetyl-muramyl--L-threonyl-D-isoglutamine (thr-MDP); N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP); N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE);  
40 RIBI, which contains three components from bacteria; monophosphoryl lipid A; trehalose dimycolate; cell wall skeleton (MPL + TDM + CWS) in a 2% squalene/Tween 80

-34-

5 emulsion; and cholera toxin. Others which may be used are non-toxic derivatives of cholera toxin, including its B subunit, and/or conjugates or genetically engineered fusions of the *H. pylori* polypeptide with cholera toxin or its B subunit, procholeraagenoid, fungal polysaccharides, including schizophyllan, muramyl dipeptide, muramyl dipeptide derivatives, phorbol esters, labile toxin of *E. coli*, non-*H. pylori* bacterial lysates, block  
10 polymers or saponins.

Other suitable delivery methods include biodegradable microcapsules or immuno-stimulating complexes (ISCOMs), cochleates, or liposomes, genetically engineered attenuated live vectors such as viruses or bacteria, and recombinant (chimeric) virus-like particles, e.g., bluetongue. The amount of adjuvant employed will depend on the type of  
15 adjuvant used. For example, when the mucosal adjuvant is cholera toxin, it is suitably used in an amount of 5 µg to 50 µg, for example 10 µg to 35 µg. When used in the form of microcapsules, the amount used will depend on the amount employed in the matrix of the microcapsule to achieve the desired dosage. The determination of this amount is within the skill of a person of ordinary skill in the art.

20 Carrier systems in humans may include enteric release capsules protecting the antigen from the acidic environment of the stomach, and including *H. pylori* polypeptide in an insoluble form as fusion proteins. Suitable carriers for the vaccines of the invention are enteric coated capsules and polylactide-glycolide microspheres. Suitable diluents are 0.2 N NaHCO<sub>3</sub> and/or saline.

25 Vaccines of the invention can be administered as a primary prophylactic agent in adults or in children, as a secondary prevention, after successful eradication of *H. pylori* in an infected host, or as a therapeutic agent with the aim to induce an immune response in a susceptible host to prevent infection by *H. pylori*. The vaccines of the invention are administered in amounts readily determined by persons of ordinary skill in the art. Thus,  
30 for adults a suitable dosage will be in the range of 10 µg to 10 g, preferably 10 µg to 100 mg, for example 50 µg to 50 mg. A suitable dosage for adults will also be in the range of 5 µg to 500 mg. Similar dosage ranges will be applicable for children. Those skilled in the art will recognize that the optimal dose may be more or less dependant upon the patient's body weight, disease, the route of administration, and other factors. Those skilled in the art  
35 will also recognize that appropriate dosage levels can be obtained based on results with known oral vaccines such as, for example, a vaccine based on an *E. coli* lysate (6 mg dose daily up to total of 540 mg) and with an enterotoxigenic *E. coli* purified antigen (4 doses of 1 mg) (Schulman et al., *J. Urol.* 150:917-921 (1993); Boedecker et al., *American Gastroenterological Assoc.* 999:A-222 (1993)). The number of doses will depend upon the  
40 disease, the formulation, and efficacy data from clinical trials. Without intending any limitation as to the course of treatment, the treatment can be administered over 3 to 8 doses



-35-

5 for a primary immunization schedule over 1 month (Boedeker, *American Gastroenterological Assoc.* 888:A-222 (1993)).

In a preferred embodiment, a vaccine composition of the invention can be based on a killed whole *E. coli* preparation with an immunogenic fragment of an *H. pylori* protein of the invention expressed on its surface or it can be based on an *E. coli* lysate, wherein the  
10 killed *E. coli* acts as a carrier or an adjuvant.

It will be apparent to those skilled in the art that some of the vaccine compositions of the invention are useful only for preventing *H. pylori* infection, some are useful only for treating *H. pylori* infection, and some are useful for both preventing and treating *H. pylori* infection. In a preferred embodiment, the vaccine composition of the invention provides  
15 protection against *H. pylori* infection by stimulating humoral and/or cell-mediated immunity against *H. pylori*. It should be understood that amelioration of any of the symptoms of *H. pylori* infection is a desirable clinical goal, including a lessening of the dosage of medication used to treat *H. pylori*-caused disease, or an increase in the production of antibodies in the serum or mucous of patients.

20

#### Determination of candidate protein antigens for antibody and vaccine development

The selection of candidate protein antigens for vaccine development were derived from the nucleotide sequence. First, all possible open reading frames (ORF's) greater than 50 nucleotides in all six reading frames were identified and translated into amino acid  
25 sequences. Second, the identified ORF's were analyzed for homology to other known exported or membrane proteins and the ORF's were also analyzed using the discriminant analysis described by Klein, et al. (Klein, P., Kanehsia, M., and DeLisi, C. (1985) *Biochimica et Biophysica Acta* 815, 468-476) for predicting exported and membrane proteins.

30 Homology searches were performed using the BLAST algorithm contained in the Wisconsin Sequence Analysis Package (Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) to compare each predicted ORF amino acid sequence with all sequences found in the current GenBank, SWISS-PROT and PIR databases. BLAST searches for local alignments between the ORF and the databank  
35 sequences and reports a probability score which indicates the probability of finding this sequence by chance in the database. ORF's with significant homology (e.g. probabilities better than  $1 \times 10^{-6}$ ) to membrane or exported proteins represent likely protein antigens for vaccine development. Possible functions are provided to some of the *H. pylori* genes as indicated in Figure 560 based on sequence homology to genes cloned in other organisms.

40 Discriminant analysis (Klein, et al. supra) was used to examine the ORF amino acid sequences using our own software. This algorithm uses the intrinsic information contained in the ORF amino acid sequence and compares it to information derived from the properties

-36-

5 of known membrane and exported proteins. This comparison predicts which proteins will be exported, membrane associated or cytoplasmic. ORF amino acid sequences identified as exported or membrane associated by this algorithm are likely protein antigens for vaccine development.

10 Infrequently it is not possible to distinguish between multiple possible nucleotides at a given position in the nucleic acid sequence. In those cases the ambiguities are denoted by an extended alphabet as follows:

These are the official IUPAC-IUB single-letter base codes

Code	Base Description	
G	Guanine	
A	Adenine	
T	Thymine	
C	Cytosine	
R	Purine	(A or G)
Y	Pyrimidine	(C or T or U)
M	Amino	(A or C)
K	Ketone	(G or T)
S	Strong interaction	(C or G)
W	Weak interaction	(A or T)
H	Not-G	(A or C or T)
B	Not-A	(C or G or T)
V	Not-T (not-U)	(A or C or G)
D	Not-C	(A or G or T)
N	Any	(A or C or G or T)

15

The amino acid translations of this invention account for the ambiguity in the nucleic acid sequence by translating the ambiguous codon as the letter "X". In all cases, the permissible amino acid residues at a position are clear from an examination of the nucleic acid sequence based on the standard genetic code.

20

#### Production of Fragments and Analogs

25 The inventor has discovered novel gene products, e.g. bacterial surface gene products, from the organism *H. pylori*. Once an example of this core structure has been provided one skilled in the art can alter the disclosed structure (of *H. pylori* genes), e.g., by producing fragments or analogs, and test the newly produced structures for activity. Examples of prior art methods which allow the production and testing of fragments and analogs are discussed below. These, or analogous methods can be used to make and screen libraries of polypeptides, e.g., libraries of random peptides or libraries of fragments or analogues of cellular proteins for the ability to bind *H. pylori* polypeptides. Such screens  
30 are useful for discovery of inhibitors of *H. pylori*.

-37-

### 5           Generation of Fragments

Fragments of a protein can be produced in several ways, e.g., recombinantly, by proteolytic digestion, or by chemical synthesis. Internal or terminal fragments of a polypeptide can be generated by removing one or more nucleotides from one end (for a terminal fragment) or both ends (for an internal fragment) of a nucleic acid which encodes the polypeptide. Expression of the mutagenized DNA produces polypeptide fragments. Digestion with "end-nibbling" endonucleases can thus generate DNA's which encode an array of fragments. DNA's which encode fragments of a protein can also be generated by random shearing, restriction digestion or a combination of the above-discussed methods.

10           Fragments can also be chemically synthesized using techniques known in the art such as conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, peptides of the present invention may be arbitrarily divided into fragments of desired length with no overlap of the fragments, or divided into overlapping fragments of a desired length.

### Production of Altered DNA and Peptide Sequences: Random Methods

20           Amino acid sequence variants of a protein can be prepared by random mutagenesis of DNA which encodes a protein or a particular domain or region of a protein. Useful methods include PCR mutagenesis and saturation mutagenesis. A library of random amino acid sequence variants can also be generated by the synthesis of a set of degenerate oligonucleotide sequences. (Methods for screening proteins in a library of variants are elsewhere herein).

### PCR Mutagenesis

In PCR mutagenesis, reduced Taq polymerase fidelity is used to introduce random mutations into a cloned fragment of DNA (Leung et al., 1989, *Technique* 1:11-15). This is a very powerful and relatively rapid method of introducing random mutations. The DNA region to be mutagenized is amplified using the polymerase chain reaction (PCR) under conditions that reduce the fidelity of DNA synthesis by Taq DNA polymerase, e.g., by using a dGTP/dATP ratio of five and adding Mn<sup>2+</sup> to the PCR reaction. The pool of amplified DNA fragments are inserted into appropriate cloning vectors to provide random mutant libraries.

### Saturation Mutagenesis

Saturation mutagenesis allows for the rapid introduction of a large number of single base substitutions into cloned DNA fragments (Mayers et al., 1985, *Science* 229:242). This technique includes generation of mutations, e.g., by chemical treatment or irradiation of single-stranded DNA *in vitro*, and synthesis of a complimentary DNA strand. The mutation frequency can be modulated by modulating the severity of the treatment, and

-38-

5 essentially all possible base substitutions can be obtained. Because this procedure does not involve a genetic selection for mutant fragments both neutral substitutions, as well as those that alter function, are obtained. The distribution of point mutations is not biased toward conserved sequence elements.

#### 10 Degenerate Oligonucleotides

A library of homologs can also be generated from a set of degenerate oligonucleotide sequences. Chemical synthesis of a degenerate sequences can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The synthesis of degenerate oligonucleotides is known in  
15 the art (see for example, Narang, SA (1983) *Tetrahedron* 39:3; Itakura et al. (1981) *Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp273-289; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477. Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et  
20 al. (1990) *Science* 249:386-390; Roberts et al. (1992) *PNAS* 89:2429-2433; Devlin et al. (1990) *Science* 249: 404-406; Cwirla et al. (1990) *PNAS* 87: 6378-6382; as well as U.S. Patents Nos. 5,223,409, 5,198,346, and 5,096,815).

#### Production of Altered DNA and Peptide Sequences: Methods for Directed Mutagenesis

25 Non-random or directed, mutagenesis techniques can be used to provide specific sequences or mutations in specific regions. These techniques can be used to create variants which include, e.g., deletions, insertions, or substitutions, of residues of the known amino acid sequence of a protein. The sites for mutation can be modified individually or in series, e.g., by (1) substituting first with conserved amino acids and then with more radical choices  
30 depending upon results achieved, (2) deleting the target residue, or (3) inserting residues of the same or a different class adjacent to the located site, or combinations of options 1-3.

#### Alanine Scanning Mutagenesis

Alanine scanning mutagenesis is a useful method for identification of certain  
35 residues or regions of the desired protein that are preferred locations or domains for mutagenesis, Cunningham and Wells (*Science* 244:1081-1085, 1989). In alanine scanning, a residue or group of target residues are identified (e.g., charged residues such as Arg, Asp, His, Lys, and Glu) and replaced by a neutral or negatively charged amino acid (most preferably alanine or polyalanine). Replacement of an amino acid can affect the interaction  
40 of the amino acids with the surrounding aqueous environment in or outside the cell. Those domains demonstrating functional sensitivity to the substitutions are then refined by introducing further or other variants at or for the sites of substitution. Thus, while the site

-39-

5 for introducing an amino acid sequence variation is predetermined, the nature of the mutation per se need not be predetermined. For example, to optimize the performance of a mutation at a given site, alanine scanning or random mutagenesis may be conducted at the target codon or region and the expressed desired protein subunit variants are screened for the optimal combination of desired activity.

10

#### Oligonucleotide-Mediated Mutagenesis

Oligonucleotide-mediated mutagenesis is a useful method for preparing substitution, deletion, and insertion variants of DNA, see, e.g., Adelman et al., (*DNA* 2:183, 1983). Briefly, the desired DNA is altered by hybridizing an oligonucleotide encoding a  
15 mutation to a DNA template, where the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or native DNA sequence of the desired protein. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the desired protein DNA. Generally,  
20 oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule. The oligonucleotides are readily synthesized using techniques known in the art such as that described by Crea et  
25 al. (*Proc. Natl. Acad. Sci. USA*, 75: 5765[1978]).

#### Cassette Mutagenesis

Another method for preparing variants, cassette mutagenesis, is based on the technique described by Wells et al. (*Gene*, 34:315[1985]). The starting material is a  
30 plasmid (or other vector) which includes the protein subunit DNA to be mutated. The codon(s) in the protein subunit DNA to be mutated are identified. There must be a unique restriction endonuclease site on each side of the identified mutation site(s). If no such restriction sites exist, they may be generated using the above-described oligonucleotide-mediated mutagenesis method to introduce them at appropriate locations in the desired  
35 protein subunit DNA. After the restriction sites have been introduced into the plasmid, the plasmid is cut at these sites to linearize it. A double-stranded oligonucleotide encoding the sequence of the DNA between the restriction sites but containing the desired mutation(s) is synthesized using standard procedures. The two strands are synthesized separately and then hybridized together using standard techniques. This double-stranded oligonucleotide is  
40 referred to as the cassette. This cassette is designed to have 3' and 5' ends that are compatible with the ends of the linearized plasmid, such that it can be directly ligated to the plasmid. This plasmid now contains the mutated desired protein subunit DNA sequence.

-40-

### 5           Combinatorial Mutagenesis

Combinatorial mutagenesis can also be used to generate mutants (Ladner et al., WO 88/06630). In this method, the amino acid sequences for a group of homologs or other related proteins are aligned, preferably to promote the highest homology possible. All of the amino acids which appear at a given position of the aligned sequences can be selected  
10       to create a degenerate set of combinatorial sequences. The variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For example, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential sequences are expressible as individual peptides, or alternatively, as a set of larger fusion  
15       proteins containing the set of degenerate sequences.

### Primary High-Through-Put Methods for Screening Libraries of Peptide Fragments or Homologs

Various techniques are known in the art for screening generated mutant gene  
20       products. Techniques for screening large gene libraries often include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the genes under conditions in which detection of a desired activity, e.g., in this case, binding to *H. pylori* polypeptide or an interacting protein, facilitates relatively easy isolation of the vector encoding the gene whose product was  
25       detected. Each of the techniques described below is amenable to high through-put analysis for screening large numbers of sequences created, e.g., by random mutagenesis techniques.

### Two Hybrid Systems

Two hybrid assays such as the system described above (as with the other screening  
30       methods described herein), can be used to identify polypeptides, e.g., fragments or analogs of a naturally-occurring *H. pylori* polypeptide, e.g., of cellular proteins, or of randomly generated polypeptides which bind to an *H. pylori* protein. (The *H. pylori* domain is used as the bait protein and the library of variants are expressed as fish fusion proteins.) In an analogous fashion, a two hybrid assay (as with the other screening methods described  
35       herein), can be used to find polypeptides which bind a *H. pylori* polypeptide.

### Display Libraries

In one approach to screening assays, the candidate peptides are displayed on the surface of a cell or viral particle, and the ability of particular cells or viral particles to bind  
40       an appropriate receptor protein via the displayed product is detected in a "panning assay". For example, the gene library can be cloned into the gene for a surface membrane protein of a bacterial cell, and the resulting fusion protein detected by panning (Ladner et al., WO 88/06630; Fuchs et al. (1991) *Bio/Technology* 9:1370-1371; and Goward et al. (1992) *TIBS*

-41-

5 18:136-140). In a similar fashion, a detectably labeled ligand can be used to score for potentially functional peptide homologs. Fluorescently labeled ligands, e.g., receptors, can be used to detect homologs which retain ligand-binding activity. The use of fluorescently labeled ligands, allows cells to be visually inspected and separated under a fluorescence microscope, or, where the morphology of the cell permits, to be separated by a  
10 fluorescence-activated cell sorter.

A gene library can be expressed as a fusion protein on the surface of a viral particle. For instance, in the filamentous phage system, foreign peptide sequences can be expressed on the surface of infectious phage, thereby conferring two significant benefits. First, since these phage can be applied to affinity matrices at concentrations well over  $10^{13}$  phage per  
15 milliliter, a large number of phage can be screened at one time. Second, since each infectious phage displays a gene product on its surface, if a particular phage is recovered from an affinity matrix in low yield, the phage can be amplified by another round of infection. The group of almost identical *E. coli* filamentous phages M13, fd., and f1 are most often used in phage display libraries. Either of the phage gIII or gVIII coat proteins  
20 can be used to generate fusion proteins without disrupting the ultimate packaging of the viral particle. Foreign epitopes can be expressed at the NH<sub>2</sub>-terminal end of pIII and phage bearing such epitopes recovered from a large excess of phage lacking this epitope (Ladner et al. PCT publication WO 90/02909; Garrard et al., PCT publication WO 92/09690; Marks et al. (1992) *J. Biol. Chem.* 267:16007-16010; Griffiths et al. (1993) *EMBO J* 12:725-734;  
25 Clackson et al. (1991) *Nature* 352:624-628; and Barbas et al. (1992) *PNAS* 89:4457-4461).

A common approach uses the maltose receptor of *E. coli* (the outer membrane protein, LamB) as a peptide fusion partner (Charbit et al. (1986) *EMBO* 5, 3029-3037). Oligonucleotides have been inserted into plasmids encoding the LamB gene to produce peptides fused into one of the extracellular loops of the protein. These peptides are  
30 available for binding to ligands, e.g., to antibodies, and can elicit an immune response when the cells are administered to animals. Other cell surface proteins, e.g., OmpA (Schorr et al. (1991) *Vaccines* 91, pp. 387-392), PhoE (Agterberg, et al. (1990) *Gene* 88, 37-45), and PAL (Fuchs et al. (1991) *Bio/Tech* 9, 1369-1372), as well as large bacterial surface structures have served as vehicles for peptide display. Peptides can be fused to pilin, a  
35 protein which polymerizes to form the pilus-a conduit for interbacterial exchange of genetic information (Thiry et al. (1989) *Appl. Environ. Microbiol.* 55, 984-993). Because of its role in interacting with other cells, the pilus provides a useful support for the presentation of peptides to the extracellular environment. Another large surface structure used for peptide display is the bacterial motile organ, the flagellum. Fusion of peptides to the  
40 subunit protein flagellin offers a dense array of many peptide copies on the host cells (Kuwajima et al. (1988) *Bio/Tech.* 6, 1080-1083). Surface proteins of other bacterial species have also served as peptide fusion partners. Examples include the *Staphylococcus*

-42-

5 protein A and the outer membrane protease IgA of *Neisseria* (Hansson et al. (1992) *J. Bacteriol.* 174, 4239-4245 and Klauser et al. (1990) *EMBO J.* 9, 1991-1999).

10 In the filamentous phage systems and the LamB system described above, the physical link between the peptide and its encoding DNA occurs by the containment of the DNA within a particle (cell or phage) that carries the peptide on its surface. Capturing the peptide captures the particle and the DNA within. An alternative scheme uses the DNA-binding protein LacI to form a link between peptide and DNA (Cull et al. (1992) *PNAS USA* 89:1865-1869). This system uses a plasmid containing the LacI gene with an oligonucleotide cloning site at its 3'-end. Under the controlled induction by arabinose, a LacI-peptide fusion protein is produced. This fusion retains the natural ability of LacI to  
15 bind to a short DNA sequence known as LacO operator (LacO). By installing two copies of LacO on the expression plasmid, the LacI-peptide fusion binds tightly to the plasmid that encoded it. Because the plasmids in each cell contain only a single oligonucleotide sequence and each cell expresses only a single peptide sequence, the peptides become specifically and stably associated with the DNA sequence that directed its synthesis. The  
20 cells of the library are gently lysed and the peptide-DNA complexes are exposed to a matrix of immobilized receptor to recover the complexes containing active peptides. The associated plasmid DNA is then reintroduced into cells for amplification and DNA sequencing to determine the identity of the peptide ligands. As a demonstration of the practical utility of the method, a large random library of dodecapeptides was made and  
25 selected on a monoclonal antibody raised against the opioid peptide dynorphin B. A cohort of peptides was recovered, all related by a consensus sequence corresponding to a six-residue portion of dynorphin B. (Cull et al. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89-1869)

This scheme, sometimes referred to as peptides-on-plasmids, differs in two important ways from the phage display methods. First, the peptides are attached to the C-terminus of the fusion protein, resulting in the display of the library members as peptides  
30 having free carboxy termini. Both of the filamentous phage coat proteins, pIII and pVIII, are anchored to the phage through their C-termini, and the guest peptides are placed into the outward-extending N-terminal domains. In some designs, the phage-displayed peptides are presented right at the amino terminus of the fusion protein. (Cwirla. et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 6378-6382) A second difference is the set of biological biases affecting the population of peptides actually present in the libraries. The LacI fusion molecules are confined to the cytoplasm of the host cells. The phage coat fusions are exposed briefly to the cytoplasm during translation but are rapidly secreted through the inner membrane into the periplasmic compartment, remaining anchored in the membrane  
35 by their C-terminal hydrophobic domains, with the N-termini, containing the peptides, protruding into the periplasm while awaiting assembly into phage particles. The peptides in the LacI and phage libraries may differ significantly as a result of their exposure to  
40



-43-

5 different proteolytic activities. The phage coat proteins require transport across the inner membrane and signal peptidase processing as a prelude to incorporation into phage. Certain peptides exert a deleterious effect on these processes and are underrepresented in the libraries (Gallop et al. (1994) *J. Med. Chem.* 37(9):1233-1251). These particular biases are not a factor in the LacI display system.

10 The number of small peptides available in recombinant random libraries is enormous. Libraries of  $10^7$ - $10^9$  independent clones are routinely prepared. Libraries as large as  $10^{11}$  recombinants have been created, but this size approaches the practical limit for clone libraries. This limitation in library size occurs at the step of transforming the DNA containing randomized segments into the host bacterial cells. To circumvent this  
15 limitation, an *in vitro* system based on the display of nascent peptides in polysome complexes has recently been developed. This display library method has the potential of producing libraries 3-6 orders of magnitude larger than the currently available phage/phagemid or plasmid libraries. Furthermore, the construction of the libraries, expression of the peptides, and screening, is done in an entirely cell-free format.

20 In one application of this method (Gallop et al. (1994) *J. Med. Chem.* 37(9):1233-1251), a molecular DNA library encoding  $10^{12}$  decapeptides was constructed and the library expressed in an *E. coli* S30 *in vitro* coupled transcription/translation system. Conditions were chosen to stall the ribosomes on the mRNA, causing the accumulation of a substantial proportion of the RNA in polysomes and yielding complexes containing nascent  
25 peptides still linked to their encoding RNA. The polysomes are sufficiently robust to be affinity purified on immobilized receptors in much the same way as the more conventional recombinant peptide display libraries are screened. RNA from the bound complexes is recovered, converted to cDNA, and amplified by PCR to produce a template for the next round of synthesis and screening. The polysome display method can be coupled to the  
30 phage display system. Following several rounds of screening, cDNA from the enriched pool of polysomes was cloned into a phagemid vector. This vector serves as both a peptide expression vector, displaying peptides fused to the coat proteins, and as a DNA sequencing vector for peptide identification. By expressing the polysome-derived peptides on phage, one can either continue the affinity selection procedure in this format or assay the peptides  
35 on individual clones for binding activity in a phage ELISA, or for binding specificity in a completion phage ELISA (Barret, et al. (1992) *Anal. Biochem* 204:357-364). To identify the sequences of the active peptides one sequences the DNA produced by the phagemid host.

#### 40 Secondary Screens

The high through-put assays described above can be followed by secondary screens in order to identify further biological activities which will, e.g., allow one skilled in the art

-44-

5 to differentiate agonists from antagonists. The type of a secondary screen used will depend on the desired activity that needs to be tested. For example, an assay can be developed in which the ability to inhibit an interaction between a protein of interest and its respective ligand can be used to identify antagonists from a group of peptide fragments isolated though one of the primary screens described above.

10 Therefore, methods for generating fragments and analogs and testing them for activity are known in the art. Once the core sequence of interest is identified, it is routine to perform for one skilled in the art to obtain analogs and fragments.

### Peptide Mimetics

15 The invention also provides for reduction of the protein binding domains of the subject *H. pylori*-like family polypeptides, e.g., an *H. pylori* polypeptide, to generate mimetics, e.g. peptide or non-peptide agents. The peptide mimetics are able to disrupt binding of a *H. pylori* to its counter ligand, e.g., in the case of an *H. pylori* polypeptide binding to a naturally occurring ligand. The critical residues of a subject *H. pylori*  
20 polypeptide which are involved in molecular recognition of a polypeptide can be determined and used to generate *H. pylori*-derived peptidomimetics which competitively or noncompetatively inhibit binding of the *H. pylori* polypeptide with an interacting polypeptide (see, for example, "Peptide inhibitors of human papillomavirus protein binding to retinoblastoma gene protein" European patent applications EP-412,762A and EP-  
25 B31.080A). For example, scanning mutagenesis can be used to map the amino acid residues of a particular *H. pylori* polypeptide involved in binding an interacting polypeptide, peptidomimetic compounds (e.g. diazepam or isoquinoline derivatives) can be generated which mimic those residues in binding to an interacting polypeptide, and which therefore can inhibit binding of an *H. pylori* polypeptide to an interacting polypeptide and  
30 thereby interfere with the function of *H. pylori* polypeptide. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988),  
35 substituted gamma lactam rings (Garvey et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudo-peptides (Ewenson et al. (1986) *J Med Chem* 29:295; and Ewenson et al. in *Peptides: Structure and Function* (Proceedings of the 9th American Peptide Symposium) Pierce Chemical Co. Rockland, IL, 1985),  $\beta$ -turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26:647; and Sato et al. (1986) *J Chem Soc Perkin Trans* 1:1231), and  $\beta$ -aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun* 126:419; and Dann et al. (1986) *Biochem Biophys Res Commun* 134:71).

-45-

5        Kits

The nucleic acid, peptides and antibodies of the present invention can be combined with other reagents and articles to form kits. Kits for diagnostic purposes typically comprise the nucleic acid, peptides or antibodies in vials or other suitable vessels. Kits typically comprise other reagents for performing hybridization reactions, polymerase chain reactions (PCR), or for reconstitution of lyophilized components, such as aqueous media, salts, buffers, and the like. Kits may also comprise reagents for sample processing such as detergents, chaotropic salts and the like. Kits may also comprise immobilization means such as particles, supports, wells, dipsticks and the like. Kits may also comprise labeling means such as dyes, developing reagents, radioisotopes, fluorescent agents, luminescent or chemiluminescent agents, enzymes, intercalating agents and the like. With the nucleic acid and amino acid sequence information provided herein, individuals skilled in art can readily assemble kits to serve their particular purpose.

Exemplification20    1. Cloning and Sequencing of *H. pylori* DNA

*H. pylori* chromosomal DNA was isolated according to a basic DNA protocol outlined in Schleif R.F. and Wensink P.C., *Practical Methods in Molecular Biology*, p.98. Springer-Verlag, NY., 1981, with minor modifications. Briefly, cells were pelleted, resuspended in TE (10 mM Tris, 1 mM EDTA, pH 7.6) and GES lysis buffer (5.1 M guanidium thiocyanate, 0.1 M EDTA, pH 8.0, 0.5% N-laurylsarcosine) was added. Suspension was chilled and ammonium acetate (NH<sub>4</sub>Ac) was added to final concentration of 2.0 M. DNA was extracted, first with chloroform, then with phenol-chloroform, and reextracted with chloroform. DNA was precipitated with isopropanol, washed twice with 70% EtOH, dried and resuspended in TE.

Following isolation whole genomic *H. pylori* DNA was nebulized (Bodenteich et al., *Automated DNA Sequencing and Analysis* (J.C. Venter, ed.), Academic Press, 1994) to a median size of 2000 bp. After nebulization, the DNA was concentrated and separated on a standard 1% agarose gel. Several fractions, corresponding to approximate sizes 900-1300 bp, 1300-1700 bp, 1700-2200 bp, 2200-2700 bp, were excised from the gel and purified by the GeneClean procedure (Bio101, Inc.).

The purified DNA fragments were then blunt-ended using T4 DNA polymerase. The healed DNA was then ligated to unique BstXI-linker adapters in 100-1000 fold molar excess. These linkers are complimentary to the BstXI-cut pMPX vectors, while the overhang is not self-complimentary. Therefore, the linkers will not concatemerize nor will the cut-vector religate itself easily. The linker-adopted inserts were separated from the unincorporated linkers on a 1% agarose gel and purified using GeneClean. The linker-adapted inserts were then ligated to each of the 20 pMPX vectors to construct a series of

-46-

5 "shotgun" subclone libraries. The vectors contain an out-of-frame lacZ gene at the cloning site which becomes in-frame in the event that an adapter-dimer is cloned, allowing these to be avoided by their blue-color.

All subsequent steps were based on the multiplex DNA sequencing protocols outlined in Church G.M. and Kieffer-Higgins S., *Science* 240:185-188, 1988. Only major  
10 modifications to the protocols are highlighted. Briefly, each of the 20 vectors was then transformed into DH5 $\alpha$  competent cells (Gibco/BRL, DH5 $\alpha$  transformation protocol). The libraries were assessed by plating onto antibiotic plates containing ampicillin, methicillin and IPTG/Xgal. The plates were incubated overnight at 37°C. Successful transformants were then used for plating of clones and pooling into the multiplex pools. The clones were  
15 picked and pooled into 40 ml growth medium cultures. The cultures were grown overnight at 37°C. DNA was purified using the Qiagen Midi-prep kits and Tip-100 columns (Qiagen, Inc.). In this manner, 100  $\mu$ g of DNA was obtained per pool. 15 96-well plates of DNA were generated to obtain a 5-10 fold sequence redundancy with 250-300 base average read-lengths.

20 These purified DNA samples were then sequenced using the multiplex DNA sequencing based on chemical degradation methods (Church G.M. and Kieffer-Higgins S., *Science* 240:185-188, 1988) or by Sequithrem (Epicenter Technologies) dideoxy sequencing protocols. The sequencing reactions were electrophoresed and transferred onto nylon membranes by direct transfer electrophoresis from 40 cm gels (Richterich P. and  
25 Church G.M., *Methods in Enzymology* 218:187-222, 1993) or by electroblotting (Church, *supra*). 24 samples were run per gel. 45 successful membranes were produced by chemical sequencing and 8 were produced by dideoxy sequencing. The DNA was covalently bound to the membranes by exposure to ultraviolet light, and hybridized with labeled oligonucleotides complimentary to tag sequences on the vectors (Church, *supra*). The  
30 membranes were washed to rinse off non-specifically bound probe, and exposed to X-ray film to visualize individual sequence ladders. After autoradiography, the hybridized probe was removed by incubation at 65°C, and the hybridization cycle repeated with another tag sequence until the membrane has been probed 38 times for chemical sequencing membranes and 10 times for the dideoxy sequencing membranes. Thus, each gel produced  
35 a large number of films, each containing new sequencing information. Whenever a new blot was processed, it was initially probed for an internal standard sequence added to each of the pools.

Digital images of the films were generated using a laser-scanning densitometer (Molecular Dynamics, Sunnyvale, CA). The digitized images were processed on computer  
40 workstations (VaxStation 4000's) using the program REPLICA™ (Church et al., *Automated DNA Sequencing and Analysis* (J.C. Venter, ed.), Academic Press, 1994).

-47-

5 Image processing included lane straightening, contrast adjustment to smooth out intensity differences, and resolution enhancement by iterative gaussian deconvolution. The sequences were then automatically picked in REPLICA™ and displayed for interactive proofreading before being stored in a project database. The proofreading was accomplished by a quick visual scan of the film image followed by mouse clicks on the bands of the  
10 displayed image to modify the base calls. For typical sequences derived by chemical sequencing, the error rate of the REPLICA™ base calling software was 2-5% with most errors occurring near the end of a sequence read. Many of the sequence errors could be detected and corrected because multiple sequence reads covering the same portion of the genomic DNA provide adequate sequence redundancy for editing. Each sequence  
15 automatically received a number correspond to (microtiter plate and probe information) and lane set number (corresponding to microtiter plate columns). This number serves as a permanent identifier of the sequence so it is always possible to identify the original of any particular sequence without recourse to a specialized database.

Routine assembly of *H. pylori* sequences was done using the program FALCON  
20 (Church, Church et al., *Automated DNA Sequencing and Analysis* (J.C. Venter, ed.), Academic Press, 1994). This program has proven to be fast and reliable for most sequences. The assembled contigs were displayed using a modified version of GelAssemble, developed by the Genetics Computer Group (GCG) (Devereux et al., *Nucleic Acid Res.* 12:387-95, 1984) that interacts with REPLICA™. This provided for an  
25 integrated editor that allows multiple sequence gel images to be instantaneously called up from the REPLICA™ database and displayed to allow rapid scanning of contigs and proofreading of gel traces where discrepancies occurred between different sequence reads in the assembly.

## 30 II. Identification, cloning and expression of recombinant *H. pylori* DNA sequences

To facilitate the cloning, expression and purification of membrane and secreted proteins from *H. pylori* a powerful gene expression system, the pET System (Novagen), for cloning and expression of recombinant proteins in *E. coli*, was selected. Also, a DNA sequence encoding a peptide tag, the His-Tag, was fused to the 3' end of DNA sequences of  
35 interest in order to facilitate purification of the recombinant protein products. The 3' end was selected for fusion in order to avoid alteration of any 5' terminal signal sequence. The exception to the above was *ppiB*, a gene cloned for use as a control in the expression studies. In this study, the sequence for *H. pylori ppiB* contains a DNA sequence encoding a His-Tag fused to the 5' end of the full length gene, because the protein product of this gene  
40 does not contain a signal sequence and is expressed as a cytosolic protein.

-48-

- 5 *PCR Amplification and cloning of DNA sequences containing ORF's for membrane and secreted proteins from the J99 Strain of Helicobacter pylori.*

Sequences chosen (from the list of the DNA sequences of the invention) for cloning from the J99 strain of *H. pylori* were prepared for amplification cloning by polymerase chain reaction (PCR). Synthetic oligonucleotide primers (Table 2) specific for the 5' and 3' ends of open reading frames (ORFs) were designed and purchased (GibcoBRL Life Technologies, Gaithersburg, MD, USA). All forward primers (specific for the 5' end of the sequence) were designed to include an NcoI cloning site at the extreme 5' terminus, except for *H. pylori* sequence 4821082 (SEQ ID NO: 212) where NdeI was used. These primers were designed to permit initiation of protein translation at a methionine residue followed by a valine residue and the coding sequence for the remainder of the native *H. pylori* DNA sequence. An exception is *H. pylori* sequence 4821082 (SEQ ID NO: 212) where the initiator methionine is immediately followed by the remainder of the native *H. pylori* DNA sequence. All reverse primers (specific for the 3' end of any *H. pylori* ORF) included a EcoRI site at the extreme 5' terminus to permit cloning of each *H. pylori* sequence into the reading frame of the pET-28b. The pET-28b vector provides sequence encoding an additional 20 carboxy-terminal amino acids including six histidine residues (at the extreme C-terminus), which comprise the His-Tag. An exception to the above, as noted earlier, is the vector construction for the *ppiB* gene. A synthetic oligonucleotide primer specific for the 5' end of *ppiB* gene encoded a BamHI site at its extreme 5' terminus and the primer for the 3' end of the *ppiB* gene encoded a XhoI site at its extreme 5' terminus.

TABLE 2

Oligonucleotide primers used for PCR amplification of *H. pylori* DNA sequences

Outer membrane Proteins	Forward primer 5' to 3'	Reverse Primer 5' to 3'
7116626 (SEQ ID NO: 223)	5'-ATATCCATGGTGAGT TTGATGA-3' (SEQ ID NO:229)	5'-ATGAATTCAATTTT TATTTTGCCA-3' (SEQ ID NO:230)
29479681 (SEQ ID NO: 179)	5'-AATTCCATGGTGGGG GCTATG-3' (SEQ ID NO:231)	5'-ATGAATTCTCGATAG CCAAAATC-3' (SEQ ID NO:232)
<b>Periplasmic/ Secreted Proteins</b>		
30100332 (SEQ ID NO: 181)	5'-ATTTCATGGTCATG TTCATATT-3' (SEQ ID NO:233)	5'-ATGAATTCCATCTTT TATTCAC-3' (SEQ ID NO:234)
4721061 (SEQ ID NO: 211)	5'-AACCATGGTGATT TAAGCATTGAAAG-3' (SEQ ID NO:235)	5'-AAGAATTCCACTCA AAATTTTAAACAG-3' (SEQ ID NO:236)

5

Other Surface Proteins		
4821082 (SEQ ID NO: 212)	5'-GATCATCCATATGTT ATCTTCTAAT-3' (SEQ ID NO:237)	5'-TGAATTCAACCATTT TAACCCTG-3' (SEQ ID NO:238)
978477 (SEQ ID NO: 228)	5'-TATACCATGGTGAA ATTTTTCTTTTA-3' (SEQ ID NO:239)	5'-AGAATTCAATTGCG TCTTGTAAG-3' (SEQ ID NO:240)
Cytoplasmic Protein		
ppiB	5'-TTATGGATCCAAAC CAATTAATACT-3' (SEQ ID NO:241)	5'-TATCTCGAGTTATA GAGAAGGGC-3' (SEQ ID NO:242)

Genomic DNA prepared from the J99 strain of *H. pylori* (ATCC #55679) was used as the source of template DNA for PCR amplification reactions (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). To amplify a DNA sequence containing an *H. pylori* ORF, genomic DNA (50 nanograms) was introduced into a reaction vial containing 2 mM MgCl<sub>2</sub>, 1 micromolar synthetic oligonucleotide primers (forward and reverse primers) complementary to and flanking a defined *H. pylori* ORF, 0.2 mM of each deoxynucleotide triphosphate; dATP, dGTP, dCTP, dTTP and 2.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA) in a final volume of 100 microliters. The following thermal cycling conditions were used to obtain amplified DNA products for each ORF using a Perkin Elmer Cetus/ GeneAmp PCR System 9600 thermal cycler:

**Sequences 7116626 (SEQ ID NO: 223), 29479681 (SEQ ID NO: 179), 30100332 (SEQ ID NO: 181), 4821082 (SEQ ID NO: 212) and 978477 (SEQ ID NO: 228);**  
 Denaturation at 94°C for 2 min,  
 2 cycles at 94°C for 15 sec, 30°C for 15 sec and 72°C for 1.5 min  
 23 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min  
 Reactions were concluded at 72°C for 6 minutes.

**Sequence 4721061 (SEQ ID NO: 211);**  
 Denaturation at 94°C for 2 min,  
 2 cycles at 94°C for 15 sec, 36°C for 15 sec and 72°C for 1.5 min  
 23 cycles at 94°C for 15 sec, 60°C for 15 sec and 72°C for 1.5 min  
 Reactions were concluded at 72°C for 6 minutes.

**Conditions for amplification of *H. pylori* ppiB;**  
 Denaturation at 94°C for 2 min,  
 2 cycles at 94°C for 15 sec, 32°C for 15 sec and 72°C for 1.5 min  
 25 cycles at 94°C for 15 sec, 56°C for 15 sec and 72°C for 1.5 min  
 Reactions were concluded at 72°C for 6 minutes

-50-

5        Upon completion of thermal cycling reactions, each sample of amplified DNA was washed and purified using the Qiaquick Spin PCR purification kit (Qiagen, Gaithersburg, MD. USA). All amplified DNA samples were subjected to digestion with the restriction endonucleases, NcoI and EcoRI (New England BioLabs, Beverly, MA, USA), or in the case of *H. pylori* sequence 4821082 (SEQ ID NO: 212), with NdeI and EcoRI (Current  
10   Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). DNA samples were then subjected to electrophoresis on 1.0 % NuSeive (FMC BioProducts, Rockland, ME USA) agarose gels. DNA was visualized by exposure to ethidium bromide and long wave uv irradiation. DNA contained in slices isolated from the agarose gel was purified using the Bio 101 GeneClean Kit protocol (Bio 101 Vista, CA,  
15   USA).

*Cloning of H. pylori DNA sequences into the pET-28b prokaryotic expression vector.*

      The pET-28b vector was prepared for cloning by digestion with NcoI and EcoRI, or in the case of *H. pylori* sequence 4821082 (SEQ ID NO: 212) with NdeI and EcoRI  
20   (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994). In the case of cloning ppiB, the pET-28a vector, which encodes a His-Tag that can be fused to the 5' end of an inserted gene, was used and the cloning site prepared for cloning with the ppiB gene by digestion with BamHI and XhoI restriction endonucleases.

      Following digestion, DNA inserts were cloned (Current Protocols in Molecular  
25   Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994) into the previously digested pET-28b expression vector, except for the amplified insert for ppiB, which was cloned into the pET-28a expression vector. Products of the ligation reaction were then used to transform the BL21 strain of *E. coli* (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994) as described below.

30

*Transformation of competent bacteria with recombinant plasmids*

      Competent bacteria, *E. coli* strain BL21 or *E. coli* strain BL21(DE3), were transformed with recombinant pET expression plasmids carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular, John Wiley and  
35   Sons, Inc., F. Ausubel et al., eds., 1994). Briefly, 1 microliter of ligation reaction was mixed with 50 microliters of electrocompetent cells and subjected to a high voltage pulse, after which, samples were incubated in 0.45 milliliters SOC medium (0.5% yeast extract, 2.0 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub> and 20, mM glucose) at 37°C with shaking for 1 hour. Samples were then spread on LB agar plates  
40   containing 25 microgram/ml kanamycin sulfate for growth overnight. Transformed colonies of BL21 were then picked and analyzed to evaluate cloned inserts as described below.



-51-

5     *Identification of recombinant pET expression plasmids carrying H. pylori sequences*

Individual BL21 clones transformed with recombinant pET-28b-H.pylori ORFs were analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers, specific for each *H. pylori* sequence, that were used in the original PCR amplification cloning reactions. Successful amplification verified the integration of the *H.*  
10     *pylori* sequences in the expression vector (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., eds., 1994).

*Isolation and Preparation of plasmid DNA from BL21 transformants*

Individual clones of recombinant pET-28b vectors carrying properly cloned *H.*  
15     *pylori* ORFs were picked and incubated in 5 mls of LB broth plus 25 microgram/ml kanamycin sulfate overnight. The following day plasmid DNA was isolated and purified using the Qiagen plasmid purification protocol (Qiagen Inc., Chatsworth, CA, USA).

*Expression of recombinant H. pylori sequences in E. coli*

20     The pET vector can be propagated in any *E. coli* K-12 strain e.g. HMS174, HB101, JM109, DH5, etc. for the purpose of cloning or plasmid preparation. Hosts for expression include *E. coli* strains containing a chromosomal copy of the gene for T7 RNA polymerase. These hosts are lysogens of bacteriophage DE3, a lambda derivative that carries the lacI gene, the lacUV5 promoter and the gene for T7 RNA polymerase. T7 RNA polymerase is  
25     induced by addition of isopropyl-B-D-thiogalactoside (IPTG), and the T7 RNA polymerase transcribes any target plasmid, such as pET-28b, carrying a T7 promoter and a gene of interest. Strains used include: BL21(DE3) (Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Meth. Enzymol. 185, 60-89).

To express recombinant *H. pylori* sequences, 50 nanograms of plasmid DNA  
30     isolated as described above was used to transform competent BL21(DE3) bacteria as described above (provided by Novagen as part of the pET expression system kit). The lacZ gene (beta-galactosidase) was expressed in the pET-System as described for the *H. pylori* recombinant constructions. Transformed cells were cultured in SOC medium for 1 hour, and the culture was then plated on LB plates containing 25 micrograms/ml kanamycin  
35     sulfate. The following day, bacterial colonies were pooled and grown in LB medium containing kanamycin sulfate (25 micrograms/ml) to an optical density at 600 nm of 0.5 to 1.0 O.D. units, at which point, 1 millimolar IPTG was added to the culture for 3 hours to induce gene expression of the *H. pylori* recombinant DNA constructions .

After induction of gene expression with IPTG, bacteria were pelleted by  
40     centrifugation in a Sorvall RC-3B centrifuge at 3500 x g for 15 minutes at 4°C. Pellets were resuspended in 50 milliliters of cold 10 mM Tris-HCl, pH 8.0, 0.1 M NaCl and 0.1

-52-

- 5 mM EDTA (STE buffer). Cells were then centrifuged at 2000 x g for 20 min at 4°C. Wet pellets were weighed and frozen at -80°C until ready for protein purification.

### III. Purification of recombinant proteins from *E. coli*

#### *Analytical Methods*

- 10 The concentrations of purified protein preparations were quantified spectrophotometrically using absorbance coefficients calculated from amino acid content (Perkins, S.J. 1986 Eur. J. Biochem. 157, 169-180). Protein concentrations were also measured by the method of Bradford, M.M. (1976) Anal. Biochem. 72, 248-254, and Lowry, O.H., Rosebrough, N., Farr, A.L. & Randall, R.J. (1951) J. Biol. Chem. 193, pages 15 265-275, using bovine serum albumin as a standard.

- SDS-polyacrylamide gels (12% or 4.0 to 25 % acrylamide gradient gels) were purchased from BioRad (Hercules, CA, USA), and stained with Coomassie blue. Molecular weight markers included rabbit skeletal muscle myosin (200 kDa), *E. coli* (-galactosidase (116 kDa), rabbit muscle phosphorylase B (97.4 kDa), bovine serum albumin 20 (66.2 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa), egg white lysozyme (14.4 kDa) and bovine aprotinin (6.5 kDa).

#### *1. Purification of soluble proteins*

- All steps were carried out at 4°C. Frozen cells were thawed, resuspended in 5 25 volumes of lysis buffer (20 mM Tris, pH 7.9, 0.5 M NaCl, 5 mM imidazole with 10% glycerol, 0.1 % 2-mercaptoethanol, 200 µg/ml lysozyme, 1 mM phenylmethylsulfonyl fluoride (PMSF), and 10 µg/ml each of leupeptin, aprotinin, pepstatin, L-1-chloro-3-[4-tosylamido]-7-amino-2-heptanone (TLCK), L-1-chloro-3-[4-tosylamido]-4-phenyl-2-butanone (TPCK), and soybean trypsin inhibitor, and ruptured by several passages through 30 a small volume microfluidizer (Model M-110S, Microfluidics International Corporation, Newton, MA). The resultant homogenate was made 0.1 % Brij 35, and centrifuged at 100,000 x g for 1 hour to yield a clear supernatant (crude extract).

- Following filtration through a 0.8 µm Supor filter (Gelman Sciences, FRG) the crude extract was loaded directly onto a Ni<sup>2+</sup>-nitrilotriacetate-agarose (NTA) with a 5 35 milliliter bed volume (Hochuli, E., Dbeli, H., and Schacheer, A. (1987) J. Chromatography 411, 177-184) pre-equilibrated in lysis buffer containing 10 % glycerol, 0.1 % Brij 35 and 1 mM PMSF. The column was washed with 250 ml (50 bed volumes) of lysis buffer containing 10 % glycerol, 0.1 % Brij 35, and was eluted with sequential steps of lysis buffer containing 10 % glycerol, 0.05 % Brij 35, 1 mM PMSF, and 20, 100, 200, and 500 40 mM imidazole in succession. Fractions were monitored by absorbance at OD<sub>280</sub> nm. and peak fractions were analyzed by SDS-PAGE. Fractions containing the recombinant protein eluted at 100 mM imidazole.

-53-

5 *Recombinant beta-galactosidase (lacZ) and peptidyl-prolyl cis-trans isomerase (ppiB)*

Fractions containing the recombinant proteins from the  $\text{Ni}^{2+}$ -NTA-agarose columns were pooled and then concentrated to approximately 5 ml by centrifugal filtration (Centriprep-10, Amicon, MA), and loaded directly onto a 180-ml column (1.6 X 91 cm) of Sephacryl S-100 HR gel filtration medium equilibrated in Buffer A (10 mM Hepes, pH 7.5, 10 150 mM NaCl, 0.1 mM EGTA) and run in Buffer A at 18 ml/h. Fractions containing the recombinant protein were identified by absorbance at 280 nm and analyzed by SDS-PAGE. Fractions were pooled and concentrated by centrifugal filtration.

*Recombinant protein 7116626 (SEQ ID NO: 223)*

15 Fractions containing the recombinant protein from the  $\text{Ni}^{2+}$ -NTA-agarose column were pooled and dialyzed overnight against 1 liter of dialysis buffer (10 mM MOPS, pH 6.5, 50 mM NaCl, 0.1 mM EGTA, 0.02% Brij 35 and 1 mM PMSF). In the morning, a fine white precipitate was removed by centrifugation and the resulting supernatant was loaded onto an 8 ml (8 x 75 mm) MonoS high performance liquid chromatography column 20 (Pharmacia Biotechnology, Inc., Piscataway, NJ, USA) equilibrated in buffer B (10 mM MOPS, pH 6.5, 0.1 mM EGTA) containing 50 mM NaCl. The column was washed with 10 bed volumes of buffer B containing 50 mM NaCl, and developed with a 50-ml linear gradient of increasing NaCl (50 to 500 mM). Recombinant protein 7116626 (SEQ ID NO: 223) eluted as a sharp peak at 300 mM NaCl.

25

*2. Purification of insoluble proteins from inclusion bodies*

The following steps were carried out at 4°C. Cell pellets were resuspended in lysis buffer with 10% glycerol 200 µg/ml lysozyme, 5 mM EDTA, 1mM PMSF and 0.1 % -mercaptoethanol. After passage through the cell disrupter, the resulting homogenate was 30 made 0.2 % deoxycholate, stirred 10 minutes, then centrifuged at 20,000 x g, for 30 min. The pellets were washed with lysis buffer containing 10 % glycerol, 10 mM EDTA, 1% Triton X-100, 1 mM PMSF and 0.1% -mercaptoethanol, followed by several washes with lysis buffer containing 1 M urea, 1 mM PMSF and 0.1 % 2-mercaptoethanol. The resulting white pellet was composed primarily of inclusion bodies, free of unbroken cells and 35 membranous materials..

*Recombinant proteins 30100332 (SEQ ID NO: 181), 4721061 (SEQ ID NO: 211)*

The following steps were carried out at room temperature. Purified inclusion bodies were dissolved in 20 ml 8.0 M urea in lysis buffer with 1 mM PMSF and 0.1 % 2-mercaptoethanol, and incubated at room temperature for 1 hour. Materials that did not 40 dissolve were removed by centrifugation. The clear supernatant was filtered, then loaded onto a  $\text{Ni}^{2+}$ -NTA agarose column pre-equilibrated in 8.0 M urea in Lysis Buffer. The

-54-

5 column was washed with 250 ml (50 bed volumes) of lysis buffer containing 8 M urea, 1.0 mM PMSF and 0.1 % 2-mercaptoethanol, and developed with sequential steps of lysis buffer containing 8M urea, 1 mM PMSF, 0.1 % 2-mercaptoethanol and 20, 100, 200, and 500 mM imidazole in succession. Fractions were monitored by absorbance at OD<sub>280</sub> nm, and peak fractions were analyzed by SDS-PAGE. Fractions containing the recombinant  
10 protein eluted at 100 mM imidazole.

*Recombinant proteins 29479681 (SEQ ID NO: 179), 978477 (SEQ ID NO: 228)*

The pellet containing the inclusion bodies was solubilized in buffer B containing 8 M urea, 1 mM PMSF and 0.1 % 2-mercaptoethanol, and incubated for 1 hour at room  
15 temperature. Insoluble materials were removed by centrifugation at 20,000 x g for 30 min, and the cleared supernatant was loaded onto a 15 ml ( 1.6 x 7.5 cm ) SP-Sepharose column pre-equilibrated in buffer B, 6 M urea, 1 mM PMSF, 0.1 % 2-mercaptoethanol. After washing the column with 10 bed volumes, the column was developed with a linear gradient from 0 to 500 mM NaCl.

20

*Dialysis and concentration of protein samples*

Urea was removed slowly from the protein samples by dialysis against Tris-buffered saline (TBS; 10 mM Tris pH 8.0, 150 mM NaCl) containing 0.5 % deoxycholate (DOC) with sequential reduction in urea concentration as follows; 6M, 4M, 3M, 2M, 1M,  
25 0.5 M and finally TBS without any urea. Each dialysis step was conducted for a minimum of 4 hours at room temperature.

After dialysis, samples were concentrated by pressure filtration using Amicon stirred-cells. Protein concentrations were measured using the methods of Perkins (1986 Eur. J. Biochem. 157, 169-180), Bradford ((1976) Anal. Biochem. 72, 248-254) and Lowry  
30 ((1951) J. Biol. Chem. 193, pages 265-275).

The recombinant proteins purified by the methods described above are summarized in Table 3 below.

TABLE 3

J99 Sequence Identifier	Homolog identified by Blast	Gene symbol of Homolog	Bacterial fraction used to purify recombinant proteins	Method of purification	Relative MW on SDS-PAGE gel	Final Concentration of purified protein	Composition of buffer
<b>Outer Membrane Proteins</b>							
7116626 (SEQ ID NO: 223)	P26093	e(P4)	Soluble fraction	His-Tag	29 kDa	0.8 mg/ml	A
						1.85 mg/ml	C
29479681 (SEQ ID NO: 179)	P13036	fecA	Inclusions bodies	SP-Sepharose	23 kDa	2.36 mg/ml	B

-55-

						0.5 mg/ml	B
						----	as dry pellet
				gel filtration S100 HR			
Periplasmic/Secreted Protein							
3010032 (SEQ ID NO: 181)	P23847	dppA	Inclusion bodies	His-Tag	11 kDa	2.88 mg/ml	B
4721061 (SEQ ID NO: 211)	P36175	GCP	Inclusion bodies	His-Tag	38 kDa	2.8 mg/ml	B
Other Surface Proteins							
4821082 (SEQ ID NO:212)	P08089	M protein	Inclusion bodies	His-Tag	20 kDa	1.16 mg/ml	B
978477 (SEQ ID NO: 228)	L28919	FBP54	Inclusion bodies	SP- Sephadex	44 kDa	2.56 mg/ml	B
						0.3 mg/ml	B
Control Proteins with His-Tag							
	P00722	lacZ	Soluble fraction	His-Tag	116 kDa	10 mg/ml	A
				gel filtration S200 HR			
		ppiB	Soluble fraction	His-Tag	21 kDa	4.4 mg/ml	A
				gel filtration S100 HR			

5 Buffer compositions:

A=10 mM Hepes pH 7.5, 150 mM NaCl, 0.1 mM EGTA

B= 10 mM Tris pH 8.0, 150 mM NaCl, 0.5 % DOC

C= 10 mM MOPS pH 6.5, 300 mM NaCl, 0.1 EGTA

10 IV. Analysis of *H. pylori* proteins as Vaccine candidates

To investigate the immunomodulatory effect of *H. pylori* proteins, a mouse/*H. pylori* model was used. This model mimics the human *H. pylori* infection in many respects. The focus is on the effect of oral immunization in *H. pylori* infected animals in order to test the concept of therapeutic oral immunotherapy.

15

Animals

Female SPF BALB/c mice were purchased from Bomholt Breeding center (Denmark). They were kept in ordinary makrolon cages with free supply of water and food. The animals were 4-6 weeks old at arrival.

20

Infection

After a minimum of one week of acclimatization, the animals were infected with a type 2 strain (VacA negative) of *H. pylori* (strain AH244, originally isolated from an ulcer patient). In our hands, this strain has earlier proven to be a good colonizer of the mouse stomach. The bacteria were grown overnight in Brucella broth supplemented with 10 % fetal calf serum, at 37°C in a microaerophilic atmosphere (10% CO<sub>2</sub>, 5% O<sub>2</sub>). The animals

25

-56-

- 5 were given an oral dose of omeprazole (400  $\mu\text{mol/kg}$ ) and 3-5 h after this an oral inoculation of *H. pylori* in broth (approximately  $10^8$  cfu/animal). Positive take of the infection was checked in some animals 2-3 weeks after the inoculation.

### Antigens

- 10 Recombinant *H. pylori* antigens were chosen based on their association with externally exposed *H. pylori* cell membrane. These antigens were selected from the following groups: (1.) Outer Membrane Proteins; (2.) Periplastic/Secreted proteins; (3.) Outer Surface proteins; and (4.) Inner Membrane proteins. All recombinant proteins were constructed with a hexa-HIS tag for purification reasons and the non-*Helicobacter pylori* control protein ( $\beta$ -galactosidase from *E. coli*; LacZ), was constructed in the same way.

15 All antigens were given in a soluble form, i.e. dissolved in either a HEPES buffer or in a buffer containing 0.5% Deoxycholate (DOC).

The antigens are listed in Table 4 below.

- 20 Table 4  
*Helicobacter pylori* proteins

#### Outer membrane Proteins

SEQ ID NO:179

- 25 SEQ ID NO:223

#### Periplastic/Secreted proteins

SEQ ID NO:181

SEQ ID NO:211

30

#### Other cell envelope proteins

SEQ ID NO:212

SEQ ID NO:228

- 35 **Control proteins**

$\beta$ -galactosidase (LacZ)

### Immunizations

- 40 Ten animals in each group were immunized 4 times over a 34 day period (day 1, 15, 25 and 35). Purified antigens in solution or suspension were given at a dose of 100  $\mu\text{g/mouse}$ . As an adjuvant, the animals were also given 10  $\mu\text{g/mouse}$  of Cholera toxin (CT) with each immunization. Omeprazole (400  $\mu\text{mol/kg}$ ) was given orally to the animals 3-5 h prior to immunization as a way of protecting the antigens from acid degradation. Infected control animals received HEPES buffer + CT or DOC buffer + CT. Animals were sacrificed 2-4 weeks after final immunization. A general outline of the study is shown in
- 45 Table 5 below.

-57-

5 Table 5Study outline, therapeutic immunization:

Mice were all infected with *H. pylori* strain AH244 at day 30. Proteins are listed by their Seq ID #'s.

10	<u>Substance</u>	<u>Mouse strain</u> <u>n=10</u>	<u>Dose/mouse</u>	<u>Dates for</u> <u>dosing</u>
	1. Controls, PBS	Balb/c	0.3 ml	0, 14, 24, 34
	2. Cholera toxin, 10 µg	Balb/c	0.3 ml	0, 14, 24, 34
	3. Protein 179, 100 µg + CT 10 µg	Balb/c	0.3 ml	0, 14, 24, 34
15	4. Protein 181, 100 µg + CT 10 µg	Balb/c	0.3 ml	0, 14, 24, 34
	5. Protein 211, 100 µg + CT 10 µg	Balb/c	0.3 ml	0, 14, 24, 34
	6. Protein 212, 100 µg + CT 10 µg	Balb/c	0.3 ml	0, 14, 24, 34
	7. Protein 228, 100 µg + CT 10 µg	Balb/c	0.3 ml	0, 14, 24, 34
	8. Protein 223, 100 µg + CT 10 µg	Balb/c	0.3 ml	0, 14, 24, 34

20

Analysis of infection

Mucosal infection: The mice were sacrificed by CO<sub>2</sub> and cervical dislocation. The abdomen was opened and the stomach removed. After cutting the stomach along the greater curvature, it was rinsed in saline. The mucosa from the antrum and corpus of an area of 25mm<sup>2</sup> was scraped separately with a surgical scalpel. The mucosa scraping was suspended in Brucella broth and plated onto Blood Skirrow selective plates. The plates were incubated under microaerophilic conditions for 3-5 days and the number of colonies was counted. The identity of *H. pylori* was ascertained by urease and catalase test and by direct microscopy or Gram staining.

30 The urease test was performed essentially as follows. The reagent, Urea Agar Base Concentrate, was purchased from DIFCO Laboratories, Detroit, MI (Catalog # 0284-61-3). Urea agar base concentrate was diluted 1:10 with water. 1 ml of the diluted concentrate was mixed with 100-200 µl of actively growing *H. pylori* cells. Color change to magenta indicated that cells were urease positive.

35 The catalase test was performed essentially as follows. The reagent, N,N,N',N'-Tetramethyl-p-Phenylenediamine, was purchased from Sigma, St. Louis, MO (Catalog # T3134). A solution of the reagent (1% w/v in water) was prepared. *H. pylori* cells were swabbed onto Whatman filter paper and overlaid with the 1% solution. Color change to dark blue indicated that the cells were catalase positive.

40 Serum antibodies: From all mice serum was prepared from blood drawn by heart puncture. Serum antibodies were identified by regular ELISA techniques, where the specific antigens of *Helicobacter pylori* were plated.

Mucosal antibodies: Gentle scrapings of a defined part of the corpus and of 4 cm of duodenum were performed in 50% of the mice in order to detect the presence of antibodies

-58-

5 in the mucous. The antibody titers were determined by regular ELISA technique as for serum antibodies.

Statistical analysis: Wilcoxon-Mann-Whitney sign rank test was used for determination of significant effects of the antigens on *Helicobacter pylori* colonization.  $P < 0.05$  was considered significant. Because the antrum is the major colonization site for  
10 *Helicobacter* most emphasis was put upon changes in the antral colonization.

### Results

Antibodies in sera: All antigens tested given together with CT gave rise to a measurable specific titer in serum. The highest responses were seen with SEQ ID  
15 NOs: 223, 211, and 212 (see Figure 2).

Antibodies in mucus: In the mucus scrapings, specific antibodies against all antigens tested were seen. By far the strongest response was seen with SEQ ID NOs: 181, followed by 223 (see Figure 3).

### Therapeutic immunization effects:

20 All control animals (BALB/c mice) were well colonized with *H. pylori* (strain AH244) in both antrum and corpus of the stomach. Of the antigens tested 2 proteins (SEQ ID NOs: 211 and 212) gave a good and significant reduction and/or eradication of the *H. pylori* infection. The degree of colonization of the antrum was lower following immunization with SEQ ID NOs: 228 and 223 compared to control. The effect of SEQ ID  
25 NOs: 179, and 181 did not differ from control. The control protein lacZ, i.e. the non-*H. pylori* protein, had no eradication effect and in fact had higher *Helicobacter* colonization compared to the HEPES + CT control. All data are shown in Figures 4 and 5 for proteins dissolved in HEPES and DOC respectively. Data is shown as geometric mean values.  $n=8-10$  Wilcoxon-Mann-Whitney sign rank test \* =  $p < 0.05$ ;  $x/10$  = number of mice showing  
30 eradication of *H. pylori* over the total number of mice examined.

The data presented indicate that all of the *H. pylori* associated proteins included in this study, when used as oral immunogens in conjunction with the oral adjuvant CT, resulted in stimulation of an immune response as measured by specific serum and mucosal antibodies. A majority of the proteins led to a reduction, and in some cases complete  
35 clearance of the colonization of *H. pylori* in this animal model. It should be noted that the reduction or clearance was due to heterologous protection rather than homologous protection (the polypeptides were based on the *H. pylori* J99 strain sequence and used in the therapeutic immunization studies against a different (AH244) challenge strain), indicating the vaccine potential against a wide variety of *H. pylori* strains.

40 The highest colonization in the antrum was seen in animals treated with the non-*Helicobacter* protein LacZ, indicating that the effects seen with the *Helicobacter pylori* antigens were specific.



-59-

- 5 Taken together these data strongly support the use of these *H. pylori* proteins in a pharmaceutical formulation for the use in humans to treat and/or prevent *H. pylori* infections.

V. Sequence Variance Analysis of genes in *Helicobacter pylori* strains

- 10 Three genes were cloned and sequenced from several strains of *H. pylori* to compare the DNA and deduced amino acid sequences. This information was used to determine the sequence variation between the *H. pylori* strain, J99, and other *H. pylori* strains isolated from human patients.

15 *Preparation of Chromosomal DNA.*

- Cultures of *H. pylori* strains (as listed in Table 8) were grown in BLBB (1% Tryptone, 1% Peptamin 0.1% Glucose, 0.2% Yeast Extract 0.5% Sodium Chloride, 5% Fetal Bovine Serum) to an OD<sub>600</sub> of 0.2. Cells were centrifuged in a Sorvall RC-3B at 3500 x g at 4°C for 15 minutes and the pellet resuspended in 0.95 mls of 10 mM Tris-HCl, 0.1 mM EDTA (TE). Lysozyme was added to a final concentration of 1mg/ml along with SDS to 1% and RNase A + T1 to 0.5mg/ml and 5 units/ml respectively, and incubated at 37°C for one hour. Proteinase K was then added to a final concentration of 0.4mg/ml and the sample was incubated at 55°C for more than one hour. NaCl was added to the sample to a concentration of 0.65 M, mixed carefully, and 0.15 ml of 10% CTAB in 0.7M NaCl (final is 1% CTAB/70mM NaCl) was added followed by incubation at 65°C for 20 minutes. At this point, the samples were extracted with chloroform:isoamyl alcohol, extracted with phenol, and extracted again with chloroform:isoamyl alcohol. DNA was precipitated with either EtOH (1.5 x volumes) or isopropanol (0.6 x volumes) at -70°C for 10 minutes, washed in 70% EtOH and resuspended in TE.

30

*PCR Amplification and cloning.*

- Genomic DNA prepared from twelve strains of *Helicobacter pylori* was used as the source of template DNA for PCR amplification reactions (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). To amplify a DNA sequence containing an *H. pylori* ORF, genomic DNA (10 nanograms) was introduced into a reaction vial containing 2 mM MgCl<sub>2</sub>, 1 micromolar synthetic oligonucleotide primers (forward and reverse primers, see Table 6) complementary to and flanking a defined *H. pylori* ORF, 0.2 mM of each deoxynucleotide triphosphate; dATP, dGTP, dCTP, dTTP and 0.5 units of heat stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA) in a final volume of 20 microliters in duplicate reactions.

-60-

5 Table 6Oligonucleotide primers used for PCR amplification of *H. pylori* DNA sequences.

Outer membrane Proteins	Forward primer 5' to 3'	Reverse Primer 5' to 3'
SEQ ID NO:223	5'-ATATCCATGGTGAGTTTGA TGA-3' (SEQ ID NO:243)	5'-ATGAATTCAATTTTTTATTTT GCCA-3' (SEQ ID NO:244)
SEQ ID NO:179	5'-AATTCCATGGCTATCCAAAT CCG-3' (SEQ ID NO:245)	5'-ATGAATTCGCCAAAATCGTA GTATT-3' (SEQ ID NO:246)
SEQ ID NO:199	5'-GATACCATGGAATTTATGAA AAAG-3' (SEQ ID NO:247)	5'-TGAATTCGAAAAAGTGTAGT TATAC-3' (SEQ ID NO:248)

10 The following thermal cycling conditions were used to obtain amplified DNA products for each ORF using a Perkin Elmer Cetus/ GeneAmp PCR System 9600 thermal cycler:

Sequences (by SEQ ID NO:) 223 and 199;  
 Denaturation at 94°C for 2 min,  
 2 cycles at 94°C for 15 sec, 30°C for 15 sec and 72°C for 1.5 min  
 15 23 cycles at 94°C for 15 sec, 55°C for 15 sec and 72°C for 1.5 min  
 Reactions were concluded at 72°C for 6 minutes.

Sequences (by SEQ ID NO:) 179;  
 Denaturation at 94°C for 2 min,  
 20 2 cycles at 94°C for 15 sec, 30°C for 20 sec and 72°C for 2 min  
 25 cycles at 94°C for 15 sec, 55°C for 20 sec and 72°C for 2 min  
 Reactions were concluded at 72°C for 8 minutes.

25 Upon completion of thermal cycling reactions, each pair of samples were combined and used directly for cloning into the pCR cloning vector as described below.

*Cloning of H. pylori DNA sequences into the pCR TA cloning vector.*

All amplified inserts were cloned into the pCR 2.1 (pCRII in the case of *H. pylori* sequence 223) vector by the method described in the Original TA cloning kit (Invitrogen,  
 30 San Diego, CA). Products of the ligation reaction were then used to transform the TOP10F' (INVaF' in the case of *H. pylori* sequence 223) strain of *E. coli* as described below.

*Transformation of competent bacteria with recombinant plasmids*

35 Competent bacteria, *E. coli* strain TOP10F' or *E. coli* strain INVaF' were transformed with recombinant pCR expression plasmids carrying the cloned *H. pylori* sequences according to standard methods (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). Briefly, 2 microliters of 0.5 micromolar BME was added to each vial of 50 microliters of competent cells.

-61-

5 Subsequently, 2 microliters of ligation reaction was mixed with the competent cells and incubated on ice for 30 minutes. The cells and ligation mixture were then subjected to a "heat shock" at 42°C for 30 seconds, and were subsequently placed on ice for an additional 2 minutes, after which, samples were incubated in 0.45 milliliters SOC medium (0.5% yeast extract, 2.0 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub> and 20, mM glucose) at 37°C with shaking for 1 hour. Samples were then spread on LB agar plates containing 25 microgram/ml kanamycin sulfate or 100 micrograms/ml ampicillin for growth overnight. Transformed colonies of TOP10F' or INVaF' were then picked and analyzed to evaluate cloned inserts as described below.

#### 15 *Identification of recombinant PCR plasmids carrying H. pylori sequences*

Individual TOP10F' or INVaF' clones transformed with recombinant pCR-*H. pylori* ORFs were analyzed by PCR amplification of the cloned inserts using the same forward and reverse primers, specific for each *H. pylori* sequence, that were used in the original PCR amplification cloning reactions. Successful amplification verified the integration of the *H. pylori* sequences in the cloning vector (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994).

Individual clones of recombinant pCR vectors carrying properly cloned *H. pylori* ORFs were picked for sequence analysis. Sequence analysis was performed on ABI Sequencers using standard protocols (Perkin Elmer) using vector-specific primers (as found in PCR11 or pCR2.1, Invitrogen, San Diego, CA) and sequencing primers specific to the ORF as listed in Table 7 below.

Table 7

Oligonucleotide primers used for sequencing of *H. pylori* DNA sequences.

30

Outer membrane Proteins	Forward primers 5' to 3'	Reverse Primers 5' to 3'
SEQ ID NO:223	5'-TTGAACACTTTTGATTATGCGG-3' (SEQ ID NO:249) 5'-GGATTATGCGATTGTTTTACAAG-3' (SEQ ID NO:250)	5'-GTCTTTAGCAAAAATGGCGTC-3' (SEQ ID NO:251) 5'-AATGAGCGTAAGAGAGCCTTC-3' (SEQ ID NO:252)
SEQ ID NO:179	5'-CTTATGGGGGTATTGTCA-3' (SEQ ID NO:253) 5'-AGCATGTGGGTATCCAGC-3' (SEQ ID NO:254)	5'-AGGTTGTTGCCTAAAGACT-3' (SEQ ID NO:255) 5'-CTGCCTCCACCTTTGATC-3' (SEQ ID NO:256)
SEQ ID NO:199	5'-ACCAATATCAATTGGCACT-3' (SEQ ID NO:257) 5'-ACTTGGAAGCTCTGCA-3' (SEQ ID NO:258)	5'-CTTGCTTGTCATATCTAGC-3' (SEQ ID NO:259) 5'-GTTGAAGTGTTGGTGCTA-3' (SEQ ID NO:260)

-62-

	5'-CAAGCAAGTGGTTTGGTTTTAG-3' (SEQ ID NO:261) 5'-TGGAAGAGCAAATCATTGAAG-3' (SEQ ID NO:262)	5'-GCCCATAATCAAAAAGCCCAT-3' (SEQ ID NO:263) 5'-CTAAAACCAAACCACTTGCTTGTC-3' (SEQ ID NO:264)
<b>Vector Primers</b>	5'-GTAAAACGACGGCCAG-3' (SEQ ID NO:265)	5'-CAGGAAACAGCTATGAC-3' (SEQ ID NO:266)

5

### Results

The PCR error rate for the experiments described herein was determined to be 2 base changes out of 4485 bases, which is equivalent to an estimated error rate of less than or equal to 0.04%.

10

DNA sequence analysis was performed on four different open reading frames identified as genes and amplified by PCR methods from a dozen different strains of the bacterium *Helicobacter pylori*. The deduced amino acid sequences of three of the four open reading frames that were selected for this study showed statistically significant

15

BLAST homology to defined proteins present in other bacterial species. Those ORFs included: SEQ ID NO:223, homologous to lipoprotein e (P4) present in the outer membrane of *H. influenzae*; SEQ ID NO:179, homologous to fecA, an outer membrane receptor in iron (III) dicitrate transport in *E. coli*. SEQ ID NO:199 was identified as an unknown open reading frame, because it showed low homology with sequences in the public databases.

20

To assess the extent of conservation or variance in the ORFs across various strains of *H. pylori*, changes in DNA sequence and the deduced protein sequence were compared to the DNA and deduced protein sequences found in the J99 strain of *H. pylori* (see Table 8 below). Results are presented as percent identity to the J99 strain of *H. pylori* sequenced by random shotgun cloning. To control for any variations in the J99 sequence each of the four open reading frames were cloned and sequenced again from the J99 bacterial strain and that sequence information was compared to the sequence information that had been collected from inserts cloned by random shotgun sequencing of the J99 strain.

25

-63-

5 Table 8  
Multiple Strain DNA Sequence analysis of *H. pylori* Vaccine Candidates

J99 Seq. ID #:	223	223	179	179	199	199
Length of Region Sequenced:	232 a.a.	696 nt.	182 a.a.	548 nt.	273 a.a.	819 nt.
Strain Tested	AA identity	Nuc. identity	AA identity	Nuc. identity	AA identity	Nuc. Identity
J99	100.00%	100.00%	100.00%	100.00%	99.63%	99.88%
AH244	n.d.	n.d.	99.09%	96.71%	98.90%	96.45%
AH4	97.84%	95.83%	n.d.	n.d.	97.80%	95.73%
AH5	98.28%	96.12%	98.91%	96.90%	98.53%	95.73%
AH15	97.41%	95.98%	99.82%	97.99%	99.63%	96.09%
AH61	97.84%	95.98%	99.27%	97.44%	n.d.	n.d.
5155	n.d.	n.d.	99.45%	97.08%	98.53%	95.60%
5294	98.28%	95.40%	99.64%	97.26%	97.07%	95.48%
7958	97.84%	95.40%	n.d.	n.d.	99.63%	96.46%
5640	97.41%	95.69%	99.09%	97.63%	98.53%	95.48%
AH18	98.71%	95.69%	99.64%	97.44%	100.00%	95.97%
AH24	97.84%	95.40%	99.27%	96.71%	100.00%	96.46%

n.d. = not done

#### VI. Experimental Knock-Out Protocol for the Determination of Essential *H. pylori* Genes as Potential Therapeutic Targets

- 10 Therapeutic targets are chosen from genes whose protein products appear to play key roles in essential cell pathways such as cell envelope synthesis, DNA synthesis, transcription, translation, regulation and colonization/virulence.

- 15 The protocol for the deletion of portions of *H. pylori* genes/ORFs and the insertional mutagenesis of a kanamycin-resistance cassette in order to identify genes which are essential to the cell is modified from previously published methods (Labigne-Roussel et al., 1988, J. Bacteriology 170, pp. 1704-1708; Cover et al., 1994, J. Biological Chemistry 269, pp. 10566-10573; Reyrat et al., 1995, Proc. Natl. Acad. Sci. 92, pp 8768-8772). The result is a gene "knock-out."

#### 20 *Identification and Cloning of H. pylori Gene Sequences*

- 25 The sequences of the genes or ORFs (open reading frames) selected as knock-out targets are identified from the *H. pylori* genomic sequence and used to design primers to specifically amplify the genes/ORFs. All synthetic oligonucleotide primers are designed with the aid of the OLIGO program (National Biosciences, Inc., Plymouth, MN 55447, USA), and can be purchased from Gibco/BRL Life Technologies (Gaithersburg, MD, USA). If the ORF is smaller than 800 to 1000 base pairs, flanking primers are chosen outside of the open reading frame.

-64-

5           Genomic DNA prepared from the *Helicobacter pylori* HP-J99 strain (ATCC 55679)  
is used as the source of template DNA for amplification of the ORFs by PCR (polymerase  
chain reaction) (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F.  
Ausubel et al., editors, 1994). For the preparation of genomic DNA from *H. pylori*, see  
Example I. PCR amplification is carried out by introducing 10 nanograms of genomic HP-  
10 J99 DNA into a reaction vial containing 10 mM Tris pH 8.3, 50 mM KCl, 2 mM MgCl<sub>2</sub>, 2  
microMolar synthetic oligonucleotide primers (forward=F1 and reverse=R1), 0.2 mM of  
each deoxynucleotide triphosphate (dATP, dGTP, dCTP, dTTP), and 1.25 units of heat  
stable DNA polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA)  
in a final volume of 40 microliters. The PCR is carried out with Perkin Elmer  
15 Cetus/GeneAmp PCR System 9600 thermal cyclers.

Upon completion of thermal cycling reactions, each sample of amplified DNA is  
visualized on a 2% TAE agarose gel stained with Ethidium Bromide (Current Protocols in  
Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994) to  
determine that a single product of the expected size had resulted from the reaction.  
20 Amplified DNA is then washed and purified using the Qiaquick Spin PCR purification kit  
(Qiagen, Gaithersburg, MD, USA).

PCR products are cloned into the pT7Blue T-Vector (catalog#69820-1, Novagen,  
Inc., Madison, WI, USA) using the TA cloning strategy (Current Protocols in Molecular  
Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). The ligation of the  
25 PCR product into the vector is accomplished by mixing a 6 fold molar excess of the PCR  
product, 10 ng of pT7Blue-T vector (Novagen), 1 microliter of T4 DNA Ligase Buffer  
(New England Biolabs, Beverly, MA, USA), and 200 units of T4 DNA Ligase (New  
England Biolabs) into a final reaction volume of 10 microliters. Ligation is allowed to  
proceed for 16 hours at 16°C.

30 Ligation products are electroporated (Current Protocols in Molecular Biology, John  
Wiley and Sons, Inc., F. Ausubel et al., editors, 1994) into electroporation-competent XL-1  
Blue or DH5- $\alpha$  *E. coli* cells (Clontech Lab., Inc. Palo Alto, CA, USA). Briefly, 1 microliter  
of ligation reaction is mixed with 40 microliters of electrocompetent cells and subjected to  
a high voltage pulse (25 microFarads, 2.5 kV, 200 ohms) after which the samples are  
35 incubated in 0.45 ml SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM  
KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub> and 20 mM glucose) at 37°C with shaking for 1 hour.  
Samples are then spread onto LB (10 g/l bacto tryptone, 5 g/l bacto yeast extract, 10 g/l  
sodium chloride) plates containing 100 microgram/ml of Ampicillin, 0.3% X-gal, and 100  
microgram/ml IPTG. These plates are incubated overnight at 37°C. Ampicillin-resistant  
40 colonies with white color are selected, grown in 5 ml of liquid LB containing 100  
microgram/ml of Ampicillin, and plasmid DNA is isolated using the Qiagen miniprep  
protocol (Qiagen, Gaithersburg, MD, USA).

-65-

5 To verify that the correct *H. pylori* DNA inserts had been cloned, these pT7Blue plasmid DNAs are used as templates for PCR amplification of the cloned inserts, using the same forward and reverse primers used for the initial amplification of the J99 *H. pylori* sequence. Recognition of the primers and a PCR product of the correct size as visualized on a 2% TAE, ethidium bromide stained agarose gel are confirmation that the correct  
10 inserts had been cloned. Two to six such verified clones are obtained for each knock-out target, and frozen at -70°C for storage. To minimize errors due to PCR, plasmid DNA from these verified clones are pooled, and used in subsequent cloning steps.

The sequences of the genes/ORFs are again used to design a second pair of primers which flank the region of *H. pylori* DNA to be either interrupted or deleted (up to 250  
15 basepairs) within the ORFs but are oriented away from each other. The pool of circular plasmid DNAs of the previously isolated clones are used as templates for this round of PCR. Since the orientation of amplification of this pair of deletion primers is away from each other, the portion of the ORF between the primers is not included in the resultant PCR product. The PCR product is a linear piece of DNA with *H. pylori* DNA at each end and  
20 the pT7Blue vector backbone between them which, in essence, results in the deletion of a portion of the ORFs. The PCR product is visualized on a 1% TAE, ethidium bromide stained agarose gel to confirm that only a single product of the correct size has been amplified.

A Kanamycin-resistance cassette (Labigne-Roussel et al., 1988 J. Bacteriology 170,  
25 1704-1708) is ligated to this PCR product by the TA cloning method used previously (Current Protocols in Molecular Biology, John Wiley and Sons, Inc., F. Ausubel et al., editors, 1994). The Kanamycin cassette containing a *Campylobacter* kanamycin resistance gene is obtained by carrying out an EcoRI digestion of the recombinant plasmid pCTB8:kan (Cover et al., 1994, J. Biological Chemistry 269, pp. 10566-10573). The proper  
30 fragment (1.4 kb) is isolated on a 1% TAE gel, and isolated using the QIAquick gel extraction kit (Qiagen, Gaithersburg, MD, USA). The fragment is end repaired using the Klenow fill-in protocol, which involved mixing 4ug of the DNA fragment, 1 microliter of dATP, dGTP, dCTP, dTTP at 0.5 mM, 2 microliter of Klenow Buffer (New England Biolabs) and 5 units of Klenow DNA Polymerase I Large (Klenow) Fragment (New  
35 England Biolabs) into a 20 microliter reaction, incubating at 30°C for 15 min, and inactivating the enzyme by heating to 75°C for 10 minutes. This blunt-ended Kanamycin cassette is then purified through a Qiaquick column (Qiagen, Gaithersburg, MD, USA) to eliminate nucleotides. The "T" overhang is then generated by mixing 5 micrograms of the blunt-ended kanamycin cassette, 10 mM Tris pH 8.3, 50 mM KCl, 2 mM MgCl<sub>2</sub>, 5 units of  
40 DNA Polymerase (Amplitaq, Roche Molecular Systems, Inc., Branchburg, NJ, USA), 20 microliters of 5 mM dTTP, in a 100 microliter reaction and incubating the reaction for 2 hours at 37°C. The "Kan-T" cassette is purified using a QIAquick column (Qiagen,

-66-

5 Gaithersburg, MD, USA). The PCR product of the deletion primers (F2 and R2) is ligated to the Kan-T cassette by mixing 10 to 25 ng of deletion primer PCR product, 50 - 75 ng Kan-T cassette DNA, 1 microliter 10x T4 DNA Ligase reaction mixture, 0.5 microliter T4 DNA Ligase (New England Biolabs, Beverly, MA, USA) in a 10 microliter reaction and incubating for 16 hours at 16°C.

10 The ligation products are transformed into XL-1 Blue or DH5- $\alpha$  *E.coli* cells by electroporation as described previously. After recovery in SOC, cells are plated onto LB plates containing 100 microgram/ml Ampicillin and grown overnight at 37°C. These plates are then replica plated onto plates containing 25 microgram/ml Kanamycin and allowed to grow overnight. Resultant colonies have both the Ampicillin resistance gene present in the  
15 pT7Blue vector, and the newly introduced Kanamycin resistance gene. Colonies are picked into LB containing 25 microgram/ml Kanamycin and plasmid DNA is isolated from the cultured cells using the Qiagen miniprep protocol (Qiagen, Gaithersburg, MD, USA).

Several tests by PCR amplification are conducted on these plasmids to verify that the Kanamycin is inserted in the *H. pylori* gene/ORF, and to determine the orientation of  
20 the insertion of the Kanamycin-resistance gene relative to the *H. pylori* gene/ORF. To verify that the Kanamycin cassette is inserted into the *H. pylori* sequence, the plasmid DNAs are used as templates for PCR amplification with the set of primers originally used to clone the *H. pylori* gene/ORFs. The correct PCR product is the size of the deleted gene/ORF but increased in size by the addition of a 1.4 kilobase Kanamycin cassette. To  
25 avoid potential polar effects of the kanamycin resistance cassette on *H. pylori* gene expression, the orientation of the Kanamycin resistance gene with respect to the knock-out gene/ORF is determined and both orientations are eventually used in *H. pylori* transformations (see below). To determine the orientation of insertion of the kanamycin resistance gene, primers are designed from the ends of the kanamycin resistance gene  
30 ("Kan-1" 5'-ATCTTACCTATCACCTCAAAT-3' (SEQ ID NO:267), and "Kan-2" 5'-AGACAGCAACATCTTTGTGAA-3' (SEQ ID NO:268)). By using each of the cloning primers in conjunction with each of the Kan primers (4 combinations of primers), the orientation of the Kanamycin cassette relative to the *H.pylori* sequence is determined. Positive clones are classified as either in the "A" orientation (the same direction of  
35 transcription is present for both the *H. pylori* gene and the Kanamycin resistance gene), or in the "B" orientation (the direction of transcription for the *H.pylori* gene is opposite to that of the Kanamycin resistance gene). Clones which share the same orientation (A or B) are pooled for subsequent experiments and independently transformed into *H. pylori*.

#### 40 Transformation of Plasmid DNA into *H. pylori* cells

Two strains of *H. pylori* are used for transformation: HP-J99 (ATCC 55679), the clinical isolate which provided the DNA from which the *H. pylori* sequence database is



-67-

5 obtained, and AH244, an isolate which had been passaged in, and has the ability to colonize the mouse stomach. Cells for transformation are grown at 37°C, 10% CO<sub>2</sub>, 100% humidity, either on Sheep-Blood agar plates or in Brucella Broth liquid. Cells are grown to exponential phase, and examined microscopically to determine that the cells are "healthy" (actively moving cells) and not contaminated. If grown on plates, cells are harvested by  
10 scraping cells from the plate with a sterile loop, suspended in 1 ml of Brucella Broth, spun down (1 minute, top speed in eppendorf microfuge) and resuspended in 200 microliters Brucella Broth. If grown in Brucella Broth liquid, cells are centrifuged (15 minutes at 3000 rpm in a Beckman TJ6 centrifuge) and the cell pellet resuspended in 200 microliters of Brucella broth. An aliquot of cells is taken to determine the optical density at 600 nm, in  
15 order to calculate the concentration of cells. An aliquot (1 to 5 OD<sub>600</sub> units/25 microliter) of the resuspended cells is placed onto a prewarmed Sheep-Blood agar plate, and the plate is further incubated at 37°C, 6% CO<sub>2</sub>, 100% humidity for 4 hours. After this incubation, 10 microliters of plasmid DNA (100 micrograms per microliter) is spotted onto these cells. A positive control (plasmid DNA with the ribonuclease H gene disrupted by kanamycin  
20 resistance gene) and a negative control (no plasmid DNA) are done in parallel. The plates are returned to 37°C, 6% CO<sub>2</sub> for an additional 4 hours of incubation. Cells are then spread onto that plate using a swab wetted in Brucella broth, and grown for 20 hours at 37°C, 6% CO<sub>2</sub>. Cells are then transferred to a Sheep-Blood agar plate containing 25 micrograms/ml Kanamycin, and allowed to grow for 3 to 5 days at 37°C, 6% CO<sub>2</sub>, 100%  
25 humidity. If colonies appear, they are picked and regrown as patches on a fresh Sheep-Blood agar plate containing 25 micrograms/ml Kanamycin.

Three sets of PCR tests are done to verify that the colonies of transformants have arisen from homologous recombination at the proper chromosomal location. The template for PCR (DNA from the colony) is obtained by a rapid boiling DNA preparation method as  
30 follows. An aliquot of the colony (stab of the colony with a toothpick) is introduced into 100 microliters of 1% Triton X-100, 20 mM Tris, pH 8.5, and boiled for 6 minutes. An equal volume of phenol : chloroform (1:1) is added and vortexed. The mixture is microfuged for 5 minutes and the supernatant is used as DNA template for PCR with combinations of the following primers to verify homologous recombination at the proper  
35 chromosomal location.

TEST 1. PCR with cloning primers originally used to amplify the gene/ORF. A positive result of homologous recombination at the correct chromosomal location should show a single PCR product whose size is expected to be the size of the deleted gene/ORF but increased in size by the addition of a 1.4 kilobase Kanamycin cassette. A PCR product  
40 of just the size of the gene/ORF is proof that the gene had not been knocked out and that the transformant is not the result of homologous recombination at the correct chromosome location.

-68-

5 TEST 2. PCR with F3 (primer designed from sequences upstream of the gene/ORF and not present on the plasmid), and either primer Kan-1 or Kan-2 (primers designed from the ends of the kanamycin resistance gene), depending on whether the plasmid DNA used was of "A" or "B" orientation. Homologous recombination at the correct chromosomal location will result in a single PCR product of the expected size (i.e., from the location of  
10 F3 to the insertion site of kanamycin resistance gene). No PCR product or PCR product(s) of incorrect size(s) will prove that the plasmid had not integrated at the correct site and that the gene had not been knocked out.

TEST 3. PCR with R3 (primer designed from sequences downstream of the gene/ORF and not present on the plasmid) and either primer Kan-1 or Kan-2, depending on  
15 whether the plasmid DNA used was of "A" or "B" orientation. Homologous recombination at the correct chromosomal location will result in a single PCR product of the expected size (i.e., from the insertion site of kanamycin resistance gene to the downstream location of R3). Again, no PCR product or PCR product(s) of incorrect size(s) will prove that the plasmid had not integrated at the correct site and that the gene had not been knocked out.

20 Transformants showing positive results for all three tests above indicate that the gene is not essential for survival *in vitro*.

A negative result in any of the three above tests for each transformant indicates that the gene had not been disrupted, and that the gene is essential for survival *in vitro*.

In the event that no colonies result from two independent transformations while the  
25 positive control with the disrupted ribonuclease H plasmid DNA produces transformants, the plasmid DNA is further analyzed by PCR on DNA from transformant populations prior to plating for colony formation. This will verify that the plasmid can enter the cells and undergo homologous recombination at the correct site. Briefly, plasmid DNA is incubated according to the transformation protocol described above. DNA is extracted from the *H.*  
30 *pylori* cells immediately after incubation with the plasmid DNAs and the DNA is used as template for the above TEST 2 and TEST 3. Positive results in TEST 2 and TEST 3 would verify that the plasmid DNA could enter the cells and undergo homologous recombination at the correct chromosomal location. If TEST 2 and TEST 3 are positive, then failure to obtain viable transformants indicates that the gene is essential, and cells suffering a  
35 disruption in that gene are incapable of colony formation

#### VII. High-throughput drug screen assay

##### *Cloning, expression and protein purification*

40 Cloning, transformation, expression and purification of the *H. pylori* target gene and its protein product, e.g., an *H. pylori* enzyme, to be used in a high-throughput drug screen assay, is carried out essentially as described in Examples II and III above. Development and application of a screening assay for a particular *H. pylori* gene product, peptidyl-propyl *cis-trans* isomerase, is described below as a specific example.

5

*Enzymatic Assay*

The assay is essentially as described by Fisher (Fischer, G., et.al. (1984) *Biomed. Biochim. Acta* 43:1101-1111). The assay measures the *cis-trans* isomerization of the Ala-Pro bond in the test peptide N-succinyl-Ala-Ala-Pro-Phe-p-nitroanilide (Sigma # S-7388, lot # 84H5805). The assay is coupled with  $\alpha$ -chymotrypsin, where the ability of the protease to cleave the test peptide occurs only when the Ala-Pro bond is in *trans*. The conversion of the test peptide to the trans isomer in the assay is followed at 390 nm on a Beckman Model DU-650 spectrophotometer. The data are collected every second with an average scanning of time of 0.5 second. Assays are carried out in 35 mM Hepes, pH 8.0, in a final volume of 400  $\mu$ l, with 10  $\mu$ M  $\alpha$ -chymotrypsin (type 1-5 from bovine Pancreas, Sigma # C-7762, lot 23H7020) and 10 nM PPIase. To initiate the reaction, 10  $\mu$ l of the substrate ( 2 mM N-Succinyl-Ala-Ala-Pro-Phe-p-nitroanilide in DMSO) is added to 390  $\mu$ l of reaction mixture at room temperature.

20 *Enzymatic assay in crude bacterial extract.*

A 50 ml culture of *Helicobacter pylori* (strain J99) in Brucella broth is harvested at mid-log phase ( $OD_{600\text{ nm}} \sim 1$ ) and resuspended in lysis buffer with the following protease inhibitors: 1 mM PMSF, and 10  $\mu$ g/ml of each of aprotinin, leupeptin, pepstatine, TLCK, TPCK, and soybean trypsin inhibitor. The suspension is subjected to 3 cycles of freeze-thaw (15 minutes at  $-70^{\circ}\text{C}$ , then 30 minutes at room temperature), followed by sonication (three 20 second bursts). The lysate is centrifuged (12,000 g x 30 minutes) and the supernatant is assayed for enzymatic activity as described above.

Many *H. pylori* enzymes can be expressed at high levels and in an active form in *E. coli*. Such high yields of purified proteins provide for the design of various high throughput drug screening assays.

Other Embodiments

Included in the invention are: allelic variations; natural mutants; induced mutants; proteins encoded by DNA that hybridizes under high or low stringency conditions to a nucleic acid which encodes a polypeptide as shown in SEQ ID NOs:1-114 (for definitions of high and low stringency see Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1989, 6.3.1 - 6.3.6, hereby incorporated by reference); and, polypeptides specifically bound by antisera to *H. pylori* polypeptides, especially by antisera to an active site or binding domain of *H. pylori* polypeptide

The invention also includes fragments, preferably biologically active fragments, or analogs of *H. pylori* polypeptides. A biologically active fragment or analog is one having any *in vivo* or *in vitro* activity which is characteristic of the *H. pylori* polypeptides shown in

-70-

- 5 SEQ ID NOs:115-228, or of other naturally occurring *H. pylori* polypeptides, e.g., one or more of the biological activities described above. Especially preferred are fragments which exist *in vivo*, e.g., fragments which arise from post transcriptional processing or which arise from translation of alternatively spliced RNA's. Fragments include those expressed in native or endogenous cells as well as those made in expression systems, e.g., in CHO cells.
- 10 Because peptides such as *H. pylori* polypeptides often exhibit a range of physiological properties and because such properties may be attributable to different portions of the molecule, a useful *H. pylori* fragment or *H. pylori* analog is one which exhibits a biological activity in any biological assay for *H. pylori* activity. Most preferably the fragment or analog possesses 10%, preferably 40%, or at least 90% of the activity of *H. pylori*, in any *in vivo* or *in vitro* assay.

Analogues can differ from naturally occurring *H. pylori* polypeptides in amino acid sequence or in ways that do not involve sequence, or both. Non-sequence modifications include changes in acetylation, methylation, phosphorylation, carboxylation, or glycosylation.

- 20 Preferred analogs include *H. pylori* polypeptides (or biologically active fragments thereof) whose sequences differ from the wild-type sequence by one or more conservative amino acid substitutions or by one or more non-conservative amino acid substitutions, deletions, or insertions which do not abolish the biological activity of the *H. pylori* polypeptide. Conservative substitutions typically include the substitution of one amino acid for another with similar characteristics, e.g., substitutions within the following groups:
- 25 valine, glycine; glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. Other conservative substitutions can be taken from the table below.

30

TABLE 9  
CONSERVATIVE AMINO ACID REPLACEMENTS

For Amino Acid	Code	Replace with any of
Alanine	A	D-Ala, Gly, beta-Ala, L-Cys, D-Cys
Arginine	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Aspartic Acid	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamine	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Glutamic Acid	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala, Pro, D-Pro, $\beta$ -Ala Acp
Isoleucine	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Leu, D-Leu, Met, D-Met
Lysine	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val

-71-

Phenylalanine	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, cis-3,4, or 5-phenylproline
Proline	P	D-Pro, L-1-thioazolidine-4-carboxylic acid, D-or L-1-oxazolidine-4-carboxylic acid
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

5

Other analogs within the invention are those with modifications which increase peptide stability; such analogs may contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the peptide sequence. Also included are: analogs that include residues other than naturally occurring L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g.,  $\beta$  or  $\gamma$  amino acids; and cyclic analogs.

As used herein, the term "fragment", as applied to a *H. pylori* analog, will ordinarily be at least about 20 residues, more typically at least about 40 residues, preferably at least about 60 residues in length. Fragments of *H. pylori* polypeptides can be generated by methods known to those skilled in the art. The ability of a candidate fragment to exhibit a biological activity of *H. pylori* polypeptide can be assessed by methods known to those skilled in the art as described herein. Also included are *H. pylori* polypeptides containing residues that are not required for biological activity of the peptide or that result from alternative mRNA splicing or alternative protein processing events.

In order to obtain an *H. pylori* polypeptide, *H. pylori* polypeptide-encoding DNA can be introduced into an expression vector, the vector introduced into a cell suitable for expression of the desired protein, and the peptide recovered and purified, by prior art methods. Antibodies to the peptides and proteins can be made by immunizing an animal, e.g., a rabbit or mouse, and recovering anti-*H. pylori* polypeptide antibodies by prior art methods.

The nucleic acids and corresponding polypeptides of the invention were disclosed previously in the corresponding US application, U.S.S.N. 08/561,469, filed November 17, 1995 (Attorney Docket No.: GTN-001CP). The correlation between sequence identification numbers in the above-identified parent applications and sequence identification numbers provided herein is outlined in Table 10 below.

-72-

5

TABLE 10

Parent Nucleic Acid SEQ ID NO:	Relation-ship	New Nucleic Acid SEQ ID NO:	Parent Amino Acid SEQ ID NO:	Parent amino acid HPP #	Relation-ship	New Amino Acid SEQ ID NO:
881	a	1	385	2	a	115
882	a	2	390	9	a	116
883	a	3	401	20	a	117
884	a	4	407	26	a	118
885	a	5	409	28	a	119
886	a	6	410	29	a	120
887	a	7	413	34	a	121
888	a	8	431	55	a	122
889	a	9	435	60	a	123
890	a	10	442	68	a	124
891	a	11	445	71	a	125
892	a	12	449	75	a	126
893	a	13	463	89	a	127
894	a	14	464	90	a	128
895	a	15	467	94	a	129
896	a	16	470	97	a	130
897	a	17	474	101	a	131
898	a	18	476	103	a	132
899	a	19	477	104	a	133
900	a	20	480	107	a	134
901	a	21	485	114	a	135
902	a	22	487	116	a	136
903	a	23	502	133	a	137
904	a	24	507	139	a	138
905	a	25	508	140	a	139
906	a	26	511	144	a	140
907	a	27	515	148	a	141
908	a	28	517	150	a	142
909	a	29	521	154	a	143
910	a	30	526	161	a	144
911	a	31	534	170	a	145
912	a	32	538	175	a	146
913	a	33	541	178	a	147
914	a	34	545	183	a	148
915	a	35	549	187	a	149
916	a	36	551	189	a	150
917	a	37	552	190	a	151
918	a	38	557	195	a	152
919	a	39	559	197	a	153
920	a	40	561	199	a	154
921	a	41	569	209	a	155
922	a	42	571	211	a	156
923	a	43	580	220	a	157
924	a	44	584	224	a	158

-73-

925	a	45	594	234	a	159
926	a	46	595	235	a	160
927	a	47	615	256	a	161
928	a	48	616	257	a	162
929	a	49	624	265	a	163
930	a	50	626	267	a	164
931	a	51	628	269	a	165
932	a	52	632	273	a	166
933	a	53	634	275	a	167
934	a	54	637	279	a	168
935	a	55	641	283	a	169
936	a	56	644	287	a	170
937	a	57	645	288	a	171
938	a	58	646	289	a	172
939	a	59	648	291	a	173
940	a	60	652	296	a	174
941	a	61	662	307	a	175
942	a	62	671	316	a	176
943	a	63	672	317	a	177
944	a	64	675	320	a	178
945	a	65	677	322	a	179
946	a	66	684	331	a	180
947	a	67	685	332	a	181
948	a	68	686	333	a	182
949	a	69	693	343	a	183
950	a	70	703	356	a	184
951	a	71	704	357	a	185
952	a	72	709	363	a	186
953	a	73	710	364	a	187
954	a	74	711	366	a	188
955	a	75	715	371	a	189
956	a	76	723	380	a	190
957	a	77	724	381	a	191
958	a	78	731	388	a	192
959	a	79	734	391	a	193
960	a	80	745	406	a	194
961	a	81	753	415	a	195
962	a	82	754	416	a	196
963	a	83	758	420	a	197
964	a	84	760	422	a	198
965	a	85	764	426	a	199
966	a	86	765	427	a	200
967	a	87	766	428	a	201
968	a	88	774	437	a	202
969	a	89	776	439	a	203
970	a	90	778	441	a	204
971	a	91	785	448	a	205
972	a	92	788	452	a	206
973	a	93	793	457	a	207
974	a	94	795	459	a	208

-74-

975	a	95	797	461	a	209
976	a	96	806	471	a	210
977	a	97	812	477	a	211
978	a	98	820	486	a	212
979	a	99	823	489	a	213
980	a	100	827	493	a	214
981	a	101	833	499	a	215
982	a	102	834	500	a	216
983	a	103	842	509	a	217
984	a	104	852	521	a	218
985	a	105	854	523	a	219
986	a	106	858	529	a	220
987	a	107	863	536	a	221
988	a	108	864	539	a	222
989	a	109	865	540	a	223
990	a	110	867	542	a	224
991	a	111	872	548	a	225
992	a	112	877	553	a	226
993	a	113	878	554	a	227
994	a	114	880	556	a	228

- 5 a=sequences from USSN 08/561,469, filed November 17, 1995 (Attorney Docket No.:GTN-001CP).

### EQUIVALENTS

- 10 Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments and methods described herein. Such equivalents are intended to be encompassed by the scope of the following claims.

Other embodiments are within the following claims.



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT

(A) NAME: Astra Aktiebolag  
(B) STREET: S-151 85  
(C) CITY: Sodertalje  
(D) STATE:  
(E) COUNTRY: Sweden  
(F) POSTAL CODE (ZIP):

(ii) TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
RELATING TO HELICOBACTER PYLORI FOR  
DIAGNOSTICS AND THERAPEUTICS

(iii) NUMBER OF SEQUENCES: 268

## (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: CD-ROM ISO9660  
(B) COMPUTER:  
(C) OPERATING SYSTEM:  
(D) SOFTWARE:

## (v) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:

## (vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/561,469  
(B) FILING DATE: 17-NOV-1995

## (vii) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: LAHIVE & COCKFIELD  
(B) STREET: 60 State Street, Suite 510  
(C) CITY: Boston  
(D) STATE: Massachusetts  
(E) COUNTRY: USA  
(F) ZIP: 02109-1875

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Mandragouras, Amy E.  
(B) REGISTRATION NUMBER: 36,207  
(C) REFERENCE/DOCKET NUMBER: GTN-001CPPC

## (x) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617)227-7400  
(B) TELEFAX: (617)227-5941

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 519 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...519

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1

ATGAAAGGYC	CYATCCTATG	GCCGGCGTTT	TCTCAATTTA	GCGATCAAGA	TTTGAGCGAT	60
ATTGTGGCGT	ATCTCACTTC	TATTTTGCCT	AAAAATTTGA	GCGATAAGGA	AGTGTTGCGC	120
CAAAGTTGTC	AAAGGTGCCA	TAGCCTGGAT	TATGCTAAAG	ATAAGGCCTT	TAGCGATCCT	180
AAAGATTTAG	CCAATTATTT	AGGCTCTCAT	GCGCCTGATT	TGTCCATGAT	GATTAGGGCT	240
AAGGGCGAAC	ATGGCTTGAA	TGTTTTCATC	AACGATCCGC	AAAAGCTTTT	GCCTGGCACA	300
GCCATGCCTA	GAGTGGGATT	GAATGAAAAA	GCTCAAAAAC	AAGTCATTTT	TTATTTGGAA	360
AAAGCGGGCG	ATAGGAAAAA	GCATGAAAGG	AATACTTTAG	GGATTAAGAT	CATGATTTTC	420
TTTGCGGTGC	TGTCGTTCTT	GGCTTACGCT	GGAAAAGAAA	AGTTTGGAGC	GAAGTGCATT	480
AAATTTAAAA	AAGGGGGGAC	ATGGTTTTAT	GATTTTTTAA			519

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 186 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...186

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2

ATGCAAGAAT	TCAGTTTGTG	GTGCGATTTT	ATAGAAAGGG	ATTTTTTAGA	AAACGATTTT	60
TTAAAGCTCA	TCAATAAGGG	GGCTATTTGC	GGGRCGACGA	GTAACCCTAG	TTTGTTTGTG	120
GAAGCGATCA	CAAAAAGCGC	GTTTTATCAA	GATGAAATCG	CTAAAMCTCA	AAGGCAAAAA	180
AGCTAA						186

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 861 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...861
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3

TTGAASCCGA	TGAAAGTGAT	TCAAGTTTTT	TTATTTTCCA	ACCCTTTTTG	CGCGATTGTG	60
CCTAACACCG	AGCCAGAACA	TTTGGAGCAT	TATGACCACG	ATTTAGAACG	CTTTTCTTC	120
GCTTATAAAT	ATTTTTTAGA	CCATGCTCAA	AAAAGAGTGA	TCTATAAAGA	AGATCCTTTT	180
TTAAAAAACT	ATTCTAAAGA	CGCCATTGTT	TTAGAAAAAA	AAGACATTTA	TAATATCCAA	240
TACATTTTAA	AAGACGGAGA	GCCTTACACT	TCGTTTGAAT	TGAAAAATTT	GGGGGCTTTT	300
TTGGTGTGGG	GGTTAGGCCA	ACATAACGCC	ACGAATGCCA	GTTTGGCGAT	TTTAAGCGCT	360
TTAGATGAAT	TAAAATTTAGA	AGAAATTAGA	AATAATTYAT	TGAATTTTAA	AGGCATTAAA	420
AAACGCTTTG	ATATTTTGCA	AAAAACAAT	CTCATTCTCA	TTGATGATTA	CGCCACCAC	480
CCTACTGAAA	TTGGCRCCAC	TTTAAAAAGC	GCTAGGATTT	ATGCCAATTT	ATTGAATACG	540
CAAGAAAAAA	TTATAGTGAT	CTGGCAAGCG	CACAAATACT	CTCGCTTAAT	GGACAATTTA	600
GAAGAATTTA	AAAAATGTTT	TTTAGAGCAT	TGCGACAGGT	TGATCATTTT	ACCCGTTTAT	660
AGCGCGAGTG	AAGTTAAAAG	AGACATTGAT	TTGAAAGCCC	ATTTTAAAGCA	TTATAACCCC	720
ACCTTTTATAG	ACAGGGTGCG	TAAAAAGGGG	GATTTTTTTAG	AGCTGTTAGT	CAATGATAAT	780
GTGGTAGAAA	CGATTGAAA	AGGCTTTGTG	ATAGGCTTTG	GAGCGGGGGA	TATTACCTAT	840
CAGTTGAGAG	GCGAAATGTA	A				861

## (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 186 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...186
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4

TTGTTGCTTT	TCTTTCTTTT	GAARGGCGTC	GTTTTTCTT	TGGGCTTTTT	TTCCTTTTTT	60
GAAGAAGTCT	CTGGCTCTTT	TGRAGCTGTT	TCTTTGARCG	TGTTAGCGTT	AGTCATGGGG	120
TCTAGTYCTG	GGTTAGAAGA	ATTCTGTGTC	TTAGAAGAGC	TTATAAATTC	AGGGCTATCA	180
GTATAG						186

## (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 369 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...369

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5

ATGGGTTTTT	TAAAGTTTT	AAAACATGAC	GCTTTAGGGC	AAGTAGGGAA	TATTGTTATA	60
GGGAATTTCT	TAATAACGCT	CACTGTTTTA	GCGGTTTGTT	TTTCCTCTCA	AAGCGCTGAA	120
GAAACGACCA	TGCTCACCCT	AAGCTACACG	CTCTTTTSTA	TTTTAGGGGC	GTTTTTATTA	180
GTCGCAATCA	GCGTGGGAGC	GATCAAAAAAC	CTCAACGCGC	TTTTTTCTAA	AACAGGGGCTT	240
TTAAGCTTTT	CCTTACCCAT	TAGTTTGGA	TCTTTATTGC	TCCCTAAAAT	CTTGCTCCCC	300
AKGGTGTFTT	TTTATCTTCA	GTTTGTTCTG	GTTGTGGCG	AGCGTGCGTT	TGGGCTATTA	360
CCTTTTTTAA						369

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 564 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...564

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6

ATGTTAAAAA	CCCACTTAAG	CAGCGCTAGG	GGCGTTGTGG	TGTTGTCTAA	GATTTTACCG	60
GTGAACGTGG	TGTTAATGGT	GAGCGTGCGC	TTGTTTGAAA	AGGAATTAAA	ACGCAAACCT	120
TACTACATCA	TTGCGAGCGC	ACACAGCGAT	GAAGGTTTAG	AAAAATTAAA	AAAATTWGGG	180
GYTGATATGG	TGGKTTYCCC	TACAAAACCTY	ATGGCGCAGA	GAGTGAGCGC	GAATKGMTGG	240
TGYKTCCTGG	ATATGGAAAA	TATCTTAGAG	CGTTTTATCA	ACAAAAAAGA	CACGCTTTTA	300
GACTTAGAGG	AAGTGATTGT	CCCCAAAACC	AGCTGGCTTG	TGTTAAGGAA	ATTAAAAGAA	360
GCCCATTTTA	GAGAGATCGC	TAAAGCCTTT	GTGATTGGTA	TCACTCAAAA	AGATGGCAAA	420
TACATCCCCA	TGCCTGACGG	GGAAACGATT	ATTGCAAGCG	AATCCAAGCT	ATTGATGGTT	480
GGCACTTCAG	AAGGCGTTGC	GACCTGTAAG	CAACTCATTA	CTAGCCACCA	AAAACCAAAA	540
GAAGTGATT	ACATTTTATT	GTGA				564

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 582 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...582

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7

GTGGGGAGCT	TTTTATTCGT	GGGGCCTAGT	GGGGTAGGGA	AAACAGAATT	GGCTAAAGAA	60
TTGGCCTTGA	ATTTGMATTT	GCATTTTGAA	CGCTTTGACA	TGAGCGAATA	CAAAGAAGCC	120
CATAGCGTGG	CAAAGCTCAT	CGGAAGTCCT	AGCGGTTATG	TGGGGTTTGA	GCAAGGGGGG	180
TTATTGGTGA	ATGCGATTAA	AAAGCACCCG	CATTGTTTGC	TGCTTTTAGA	TGAGATAGAA	240
AAGGCCACCC	CTAATGTGTA	TGATTTGTTG	TTGCAGGTGA	TGGAKAACGC	CACTTTGAGC	300
GATAATTTAG	GCAACAAGGC	GAGTTTTAAG	CATGTGATAC	TGATTATGAC	KKCAARTGTG	360
GGGAGTAAGG	ATAAGGACAC	GCTAGGGTTT	TTTAGCACTA	AAAACGCCAA	GTATGATAGA	420
GCCGTTAAAG	AGCTTTTAAAC	CCCTGAATTC	CCATCCAGAA	TTGATGCCAT	CCTCCCCCTT	480
AACGCGCTCA	GTTTGGAGGA	TTTTGAAACG	CATTGTTTCT	GTGGAATTGG	ACGGGTTAAA	540
AGCCCTAGCA	CTAGAGCAAG	GCGTGATCTT	AAAATTCCAT	AA		582

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 909 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...909

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8

ATGGCTTTTC	AGGTCAATAC	AAATATCAAT	GCGATGAATG	CGCATGTGCA	ATCCGCACTC	60
ACTCAAAACG	CACTTAAAC	TTCATTGGAG	CGATTGAGTT	CAGGTTTAAG	GATCAATAAA	120
GCGGCTGATG	ACGCATCAGG	CATGACGGTG	GCGGATTCTT	TGCGTTCGCA	AGCGAGCAGT	180
TTGGGTCAAG	CGATTGCCAA	CACGAATGAC	GGCATGGGGA	TTATCCAGGT	TGCGGATAAG	240
GCTATGGATG	AGCAATTAAA	AATCTTAGAC	ACCGTTAAGG	TTAAAGCGAC	TCAAGCGGCT	300
CAAGATGGGC	AAACTACGGA	ATCTCGTAAA	GCGATTCAAT	CTGACATCGT	TCGTTTGATT	360
CAAGGTTTGG	ATAATATCGG	TAACACAACG	ACTTATAACG	GGCAAGCGTT	ATTGTCTGGT	420
CAATTCACCTA	ACAAAGAATT	CCAAGTAGGG	GCTTATTCTA	ACCAAAGCAT	TAAGGCTTCT	480
ATCGGCTCTA	CCACTTCCGA	TAAAATCGGT	CAGGTTTCGT	TCGCTACAGG	CGCGTTAATC	540
ACGGCTTCTG	GGGATATTAG	CTTGACTTTT	AAACAAGTGG	ATGGCGTGAA	TGATGTAAC	600
TTAGAGAGCG	TAAAAGTTTC	TAGTTCAGCA	GGCACAGGGA	TCGGCGTGTT	AGCAGAAGTG	660
ATCAATAAAA	ACTCTAACCG	AACAGGGGTT	AAAGCTTATG	CGAGCGTTAT	CACCACGAGC	720
GATGTGGCGG	TCCAGTCAGG	AAGTTTGAGT	AATTTAACCT	TAAATGGGAT	TCATTTGGGT	780
AATATCGCAG	ATATTAAGMR	AAACGACTCA	GACGGAAGGT	TAGTCACAGC	RATCAATGCG	840
GTCACCTCAG	AAACCGGTGT	GGWAGCTTAT	ACGGATCAAA	AAGGACGCTT	GAATTTGCGC	900
AGTATAGGT						909

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 486 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

80

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...486
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9

ATGTTTTTTT	AAACTTATCA	AAAATTACTG	GGCGCGAGCT	GTTTGGCGCT	GTATTTAGTG	60
GGCTGTGGGA	ATGGTGGTGG	CGGTGAATCG	CCGGTTGAGA	TGATTRCAAA	TAGCGAGGGT	120
ACGTTTCAAA	TCCACTCCAA	ACCACATAGC	ATTACTATTC	AAGCCCTCAA	CCTTAATACA	180
GGTAATTGTG	CTGTCAATTT	TGTTCCAGTA	AGTGAGACGT	TTCAAATGGG	TGTTTTAAGT	240
CAAGTTACTC	CAATCTCTAT	ACAGGATTTT	AAAGATATGG	CAAGCACTTA	TAAGATATTT	300
GATCAAAAGA	AAGGGTTGGC	AAACATAGCA	AATAAAATTT	CTCAATTAGA	GCAAAAGGGT	360
GTGATGATGA	AACCTCMAAC	CCTTAATTTT	GGAGAAAGTT	TAAAAGGCAT	TTCTCAAGGG	420
TGCAATATTA	TAGAGGCAGA	AATACAAACC	GACAAAGGCG	CTTGGACTTT	TAACTTTGAT	480
AAATAA						486

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 276 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...276
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10

GTGTGTTT	GGCTAGCTGA	TGTGATGGTG	GTTTTAAGCT	TGCACCTCAA	CCTAAACCCC	60
ACCAACCCTA	AATGGCTCAA	TAGGGACAGG	TTGGTTTTTA	GCGGAGGGCA	TGCGAGCGCG	120
TTAGTGTATA	GTTTGTTGCA	TTTGTGGGGC	TTTGATTTGA	GTTTAGACGA	TTTAAAGCGT	180
TTCAGGCAAT	TACACTCTAA	AACCCCAGGA	CACCCTGAAT	TACACCACAC	CGAAGGCATT	240
GAAATCACCA	CASCCTTTAG	GGCAAGGTTT	TGCTAA			276

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 561 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:

81

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...561

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11

ATGACAACAC	CGATGATTAT	TATTTCCCTA	GAAATGGGGT	TATCTTTAGT	TCCTATGCGA	60
CAATGTCTGG	TTTGCCAAGC	TCTGGCACGC	TCAATTTCTT	GGAACGGGTT	AGGCGGGAAT	120
GTCCGTAACA	CCAAAGTTTA	TGGTAAATTC	GCCGCTTACC	ACCATTGCA	AAAATATTTA	180
TTGATAGATT	TGATCGCTCG	TTTTAAAACG	CAAGGGGGCT	ATATCTTTAG	GTATAACACC	240
GATGATTACT	TGCCCTTAAA	CTCCACTTTC	TACATGGGGG	GSCTAACCAC	GGTGAGAGGC	300
TTTAGGAACG	GCTCAATCAC	ACCTAAAGAT	GAGTTTGGCT	TGTGGCTTGG	AGGCGATGGG	360
ATTTTTACCR	CTTCTACTGA	ATTGAGCTAT	GGGGTGTTAA	AAGCGGCTAA	AATGCGTTTA	420
GCGTGGTTTT	TTGACTTTGG	TTTCTTAACC	TTTAAMACCC	CAACTAGGGG	GAGTTTCTTC	480
TATAACGCTY	CCACCACGAC	GGCGAATTTT	AAAGATTATG	RCGTTGTAGG	GRCTGRGTTT	540
GARRGGGCGA	CTTGGAGGGC	T				561

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...315

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

ATGCAAGCGT	TAAAATCATT	GCTTGAAGTG	ATTACAAAAC	TCCAGAATCT	AGGCGGCTAT	60
TTGATGCATA	TAGCTATTTT	CATCATTTTT	ATTTGGATTG	GAGGRCTTAA	GTTTGTGCCT	120
TACGAAGCTG	AAGGGATCGC	CCCTTTGTGT	RCCAACCTCC	CTTTCTTTTC	TTTCATGTAT	180
AAATTTGAAA	AACCTGCATA	CAAACAACAC	AAAATGTCTG	AATCCCAATC	CATGCAAGAA	240
GAAATGCAAG	ATAACCCTAA	AATCGTTGAA	AACAAKAAT	GGCATAAAGA	AAACCGCACT	300
TCATTTAGTG	GCTGA					315

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 549 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...549

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

ATGCAGTTTG	AAGAAATGAA	AGAATTAGCC	CATCAAATTG	GCGTGTTTTA	CCATGTTGGC	60
GTTGATGGCA	TCGCGCTCTT	TTTGTGCTC	TTAAACGCTA	TCGTGGTGTT	ATTGTCCGTG	120
GTTTATGTCA	AAGAGCGTCG	TAAAGACTTT	GTGATTTGTC	TTTTATTGTT	AGAMGGGATC	180
TTAATGGGCG	TGTTTTCTTC	TCTTAATGTG	ATCTTTTTCT	ACGCTTTTTG	GGAAATCTCG	240
CTCTTGCCGG	TTTTATACCT	CATCGGTCGT	TTTGGCCGTA	ATAACAAAAT	CTATTCTGGC	300
ATGAAGTTTT	TCCTCTACAC	CTTTTATAGC	TCGTTGTGCA	TGCTTTTAGG	CATTTTATAC	360
ATCGGGTATG	ACTACGCCAA	TAATTACGGC	ATGATGAGTT	TTGATATTTT	AGACTGGTAT	420
CAGTTGAATT	TTTCTAGCGG	GATTAAAACC	TGGCTCTTTG	TAGCTTCTTT	AATAGGGATT	480
GCGGTTAAAA	TCCCGCTCTT	TCCCTTCACA	CATGGCTGCC	TTATGCGTAT	TCTAACGCCC	540
CCACTCTAG						549

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 351 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...351

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14

GTGAAAAAAT	ACGCTGAAGA	TTTTATCACC	AAAGATGAAG	TGAAATCCCCT	TTTAGAGCGC	60
TTGGCTAAAG	ACTATCCTAC	GATTGTAGAA	GAGAGTAAAA	AAATCCCCAC	CGGTGCGATC	120
CGCTCAGTCT	TGCAAGCCTT	GTTGCATGAA	AAAATCCCCA	TTAAGGACAT	GCTCACTATT	180
TTGGAAACGA	TTACTGATAT	TGCTCCATTG	GTTCAAAACG	ATGTGAATAT	CTTAACCGAA	240
CAAGTGAGGG	CGAGGCTTTC	YAGGGTGATC	ACCAACGCTT	TTAAATCTGA	AGACGGGCGT	300
TTGAAATTTT	TAACCTTTTC	TACCGATRGC	GAACAATTTT	TKGCTCAATA	A	351

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 720 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...720

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

ATGATGAAAA	ACAAACGCTC	TCAAAATAGC	CCTTATGTAA	CGCCTGACAA	CCCTTATCTA	60
ACGCTAGAAA	AAGCTTTAGG	GTATTCTTTT	AAAGACAAGC	GTTTATTGGA	GCAAGCCTTA	120



ACGCATAAA	T	CATGTAAGCT	CGCTTTAAAC	AATGAGCGCT	TGGAATTTT	GGGCGATGCG	180
GTGTTGGGCT	T	GGTGATAGG	GGAGCTGCTA	TACCATAAAT	TCTRTCAWTR	CGATGGGGGC	240
AAACTCTCTA	A	AATTAAGGGC	TTCTATTGTG	AGCGCGCATG	GTTTCACTAA	ATTAGCGAAA	300
GCGATTGCTT	T	TACAAGATTA	TTTGCGCGTT	TCTTCTTCTG	AAGAAATTTT	TAAGGGGAGG	360
GAAAAACCC	T	TATTCTRTC	AAGCGCTTTT	GAGGCTTTAA	TGGCTGGGGT	GTATTTAGAA	420
GCAGGGTTAG	C	TAAGGTGCG	TAAAATCATA	CAAAATTTAC	TCAATCGTGC	TTACAAGCGT	480
TTGGATTTGG	A	GCATTTGTT	TATGGATTAT	AAAACCGCTT	TGCAGGAATT	GACCCAAKCK	540
CAGTTTTCG	T	GATCCCCAC	TTACCAATTA	CTCCAAGAAA	AAGGCCCCGA	TCACCATAAA	600
GAATTTGAAA	T	GGCTCTATA	CATTCAAGAT	AAAATGTATG	CGACCGCTAA	AGGCAAGAGT	660
AAAAAAGAAG	C	CGAACAGCA	ATGCGCTTAT	CAAGCGCTTC	AAAACCTAAG	GAAGCCAAAT	720

## (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 687 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...687

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

TTGCTGGTCT	T	ACTCAATCT	AAAGAWTACG	CCGAATTTGA	TGTGGCCTTT	AGATATTATT	60
GTGGTTGTGG	C	ATGGGGTGT	ATGGGGGGTT	AATATGTTTG	GGAGCATGAG	CGTTAGAAGA	120
GAGAACTACTA	T	TTATGTGTC	TTTGTGGTAT	TACATCGCTA	CTTATGTGGG	TATAGCGGTG	180
ATGTATATCT	T	CAATAACCT	TTCTATCCCC	ACCTATTTTG	TCGCTGATAT	GGGGAGCGTT	240
TGGCATTMTA	T	TTCTATGTA	TTCAAGGCAGT	AATGATGCGC	TCATTCGAATG	GTGGTGGGGG	300
CATAATGCGG	T	CGCTTTTGT	CTTTACGAGT	GGGGTGATTG	GCACAATTTA	TTATTTCTTG	360
CCTAAAGAGA	G	CGGCCAGCC	TATCTTTTCT	TACAAACTCA	CTTTGTTTTC	TTTTTGGAGT	420
TTGATGTTTG	T	TTATATTTG	GGCGGGCGGG	CACCATTTGA	TCTATTCAC	CGTGSCTGAT	480
KGRGTACAAA	C	CCCTTTCTAG	CGYGTTTTCA	GTGGTGTGTA	TCTTGCTTTC	GYGGGGGACA	540
GCGATTAACA	T	GCTTTTAMC	GATGAGAGGC	CAATGGCACC	AGYTCAAAGA	AAGCCCTTTG	600
ATCAAGTTTT	T	AGTTTTAGC	CTCAACTTTC	TACATGCTTT	CCACGCTAGA	AGGCTCCATT	660
CAAGCCATTA	A	AAAGCGTGAA	CGCTTAG				687

## (2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 489 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...489

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17

ATGAAAGCAC	CCTCCCAAYA	GGATTTAAAA	AAAATCTTAG	GGATTGAAGA	AGTCATAATS	60
STATCCACAA	GCCCCATGGA	ATTACGATTG	GCCAATCAAA	AGCTAGGCAA	TCGTTTCATT	120
AAAACCTTAC	AAGCCATGAA	CGAATTAGAC	ATGGGTGCAT	TTTTTAACGC	TTACGCTCAA	180
ACAACCAAAG	ATCCCACCCA	TGCCACTAGC	TATGGCGTTT	TTGCGGCGAG	TTTGAATATG	240
GAATTGAAAA	AGGCTTTAAG	GCATTATCTT	TATGCGCAAA	CTTCTAACAT	GGTGATCAAC	300
TGCGTTAAAA	GCGTCCCCTT	ATCCCAAAAC	GACGGGCAAA	AAATCTTATT	GAGCTTGCAA	360
AGCCCTTTTA	ACCAGCTCAT	AGAAAAAACC	CTAGAACTAG	ACGAAAGCCA	CTTGTGCGCA	420
GCAAGCGTTC	AAAACGACAT	TAAGGCGATG	CAGCATGAGA	GTTTATACTC	GCGCCTTTAT	480
ATGTCTTGA						489

## (2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 180 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...180
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18

ATGGCGTTTA	TTCTAACCAC	AAACCTATTT	ATCAAGAGTT	TTACGAACTC	AATTCGCATA	60
ACGGGTTGTA	TTATCAGCCC	TAATGTGTTT	TTTGCTTATG	AATTTTGCGC	GTTAGGGTTT	120
AGAAAAGGGG	GTTAATTTT	GGATAATTTT	TCTAAATTCG	TGAGCCACAG	GTTGCAATAA	180

## (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 747 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...747
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19

GTGCGTTTTT	TCATTTTTTT	AATTCTCATT	TGCCCTTTAA	TATGCCCTTT	AATGAGCGCG	60
GATAGCGCTT	TACCTAGCGT	CAATCTCTCT	TTAAACGCTC	CTAGTGATCC	TAAACAACCTC	120
GTAACCACCC	TTAATGTCAT	CGCCTTACTC	ACGCTTTTGG	TTTTAGCCCC	ATCGTTGATT	180
TTAGTGATGA	CGAGTTTCAC	CCGTTTGATC	GTGGTGTTTT	CTTTTTTAAG	GACCGCTTTA	240
GGCAGCAAC	AAACCCCAAC	CACTCAAATT	CTAGTCTCGC	TCTCTTTGAT	ATTGACTTTT	300
TTTATCATGG	AACCTAGCTT	GAAAAAGGCT	TATGATACAG	GGATTAAGCC	TTATATGGAT	360
AAAAAGATTT	CTTACACCGA	AGCGTTTGAA	AAAAGCACTC	TGCCTTTCAA	GGAATTCATG	420
CTTAAAAACA	CACGAGAAAA	AGATCTAGCC	CTTTTTTTTA	GGATTAGGAA	TTTGCCTAAC	480

CCTAAAAACCC	CTGATGATGT	GAGCTTGAGC	GTTTAAATCC	CGGCATTTAT	GATAAGCGAG	540
TTGAAAACAG	CGTTTCAAAT	CGGCTTTTTA	CTCTACTTGC	CTTTTTTGGT	GATTGATATG	600
GTTATCAGCT	CTATTTTAAAT	GGCGATGGGT	ATGATGATGC	TCCCGCCTGT	AATGATTTCT	660
CTGCCTTTTA	AAATTTTGGT	GTTTATTCTG	GTGGATGGGT	TTAATTTATT	GACCGAAAAT	720
TTAGTGGCGA	GTTTTAAAT	GTTTAA				747

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 501 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...501

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20

TTGTTGGTTA	CTTTTTTGAA	TGGGTTTGAC	CCAAAAATCG	CTAATTAAAG	GAAAGCGTGC	60
AATGTTTATA	GYGTGGGGGT	GATTTATATT	GTAACCACCA	ACACGCTCAA	TATTTTAAGT	120
TGTGAGAGTT	TTGAAATTTT	AGAAAAAAGA	GAGCTGGATA	CAAGCGGCGT	TACTAAAAC	180
TCCACGCCGT	TTTTTCTAG	GGTTGAGGGC	ATTGATGCAG	GCACGCTAGG	GAAACTTTTT	240
TCAGGCAGTC	AATCTAAAAA	TTACTTCGCT	TACTATGACG	CTTTAGTGAA	AAAAGAAAAA	300
CGAAAAGAAG	TAAGGATTGA	AAAGAAAGAA	GAAAGGATTG	ATGCTAGAGA	AAATAAACGA	360
GAAATCAAGC	AAGAAGCCAT	TAAAGAGCCT	AAAAAAGCCA	ATCAAGGCAC	AGAAAACGCT	420
CCCACTTTAG	AAGAGAAAAS	CTACCAAAAR	GCAGAGCGAA	AATTTGACGC	TAAAGRAGRA	480
AGGGMATCGT	TCAAGRGATG	A				501

## (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 381 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...381

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21

ATGGAAAACA	GCACACTTTA	TATTGTTATT	GCCGGCTTAT	GGCTTGCTGT	AGGCTTTGGA	60
ATCTTTTATA	AGAAATTAGA	CATGCCCGTT	ATCATTGGCT	ACATTTGCAC	AGGAACGGTC	120
TTAGCGGCTT	TTTTTAAAT	TAATGATTTT	AATTTGTTGT	CTGATATTGG	TGAATTTGGT	180
ATCGTCTTTT	TAATGTTTAT	GATAGGCATT	GAGTTTAATT	TTGACAAGCT	CAAGTCCATC	240
AAACAAGAAG	TGCTCGTTT	TGGGCTTTTA	CAGGTTGTTT	TATGCGCTTT	AATCGCTTTT	300
TTATTGGGGT	ATTTTGTCT	GGGTCTTTCG	CCCATTTTTT	CCCTTGTTTT	AGGCATGGGG	360
CTTTCACCTC	CTTCAACCGC	C				381

## (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 51 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...51
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22

TTGTTACTCA TGCTTAATAA GCCAAAGCCT TTATTTTGTG CTCTTGGTGA A

51

## (2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1053 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...1053
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23

ATGGCATTAA	GGGTATTATT	ATTCTTTTGT	TTTTTGTTTT	TGCAAGCAGA	AGATAAGAGC	60
CAAGAATTAT	CATCTATACA	AAAACAAATG	GCTTTGGTGG	ATAAAAAACT	CGCCAAAGAC	120
GATAACGTGT	GGTTGAAAAA	ATTTGAAAAC	TATAAAATTT	ACAACCAAAT	TTATACTGAA	180
AAAGAGAGCG	TGAGGCAGGA	ATTAAGGCGC	TTAAAAACA	AAAAAAGCAA	GGATTTATTA	240
AAGATTAGCA	CCTTAGAGCA	TACCTTAAAG	GCTTTAGAGT	CCCAGCAAAA	AATGTTTGAA	300
AGCTATGGGG	TCAATCCTTT	TAAGGACTTG	ATAGAGCGCC	CCAATATCCC	CAATATCCCT	360
AATATCGCTA	ACCCTATTGC	GATCATTGAT	GGCATTTCCT	TCATCAAGAG	CATGCGTTTA	420
AAGCATGAAA	ATCTTAAAAA	TAACCAGACT	TCTTTAGGAG	AAGTTTAA	GCTTTTAGAT	480
CAAAAACACC	AGCTTTTAAA	TCAGTGGCAC	GCTTTGGATA	AAAGCGCGAA	ATTAAGCGAT	540
GAGATTTATC	AAACTCAAGC	CAAACGCTTA	GAATTGCAAG	GGGCTCAAAA	CATTCTAAAA	600
ACCACRATCG	GGATTTTCCA	AAAAGACAGC	GATGAAGCTA	TAAGCATTTG	CAAACTCTCA	660
GTTAAAAACC	AGCTTTTAA	ATTGGTTTAT	GTGTTTTTAG	CGGCCCTTTT	GAGCGTGGTG	720
TTTGCGTGGA	TTTTTAAAAAT	CATTTCAGT	AAATACATTG	AAAAATAATGA	GCGCGTCTAT	780
ACCGTGAATA	AAGCCATTAA	CTTCGTGAAT	GTGAGCGTGA	TCGKTKKAAT	CTTKCTTTTT	840
TCTTATTTAG	AAAACGTTAC	TTACTTGGA	ACGGTTTTAG	GCTTTGCGAG	CGCGGGCTTA	900
GCGATTCKGA	TGAAGGATTT	ATTCATGAGC	TTGCTCGGGT	GTTTATCAT	TTTGATTGGG	960
GGGAGCGTGC	ATGTGGGCGA	TAGGGTGCCT	ATCGCTAAGG	GGACGGATAT	TTTATTGGC	1020
GATGTGTTGG	ATACTTCTAA	TGTTGTACAT	TAA			1053

## (2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 300 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...300

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24

ATGAAAGAAC	AAGAATGGGA	TTTAAGCGCT	TTATTTGAAA	ATAAAGAAAG	CGCAGAAGAA	60
TTTTTAAAAA	CCTTACAAAC	AGAAGTGCAA	GAATTTGAGA	ACGCTTATCA	AAATAACCTT	120
AAGAATTTAG	ACGCTGCAAA	ATTTGCCAAC	ACTCTTAAAC	ATTACGAAAA	TTTGTCAGAA	180
AAGATCTCTA	GAGCGATGGC	TTACGCCAAT	TACTTTTTGC	CAAGAACACT	AAAGAAGCGA	240
AGTTTTATTC	GCAATGCAAA	TGGCTTGTGC	AAATATCCAA	CAACACCTTT	TATTCTTTGA	300

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 237 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...237

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25

TTGCGCGTGG	GCATGTATGA	AGTGTGTAAC	CATAAAGACG	GCACCGCTTA	TCATTCCACA	60
AGAGGTTCTA	AGGTTACCTT	AGCGTGTAAC	ACCGGCACCG	CGCAAGTCGT	AGAAATCGCT	120
CAAAACATCG	TCAATCGCAT	GAAAGAAAAG	GATATGGAAT	ATTTCATCS	MTCCCATRCG	180
TGGATTACGR	CATATCTTTR	CCCTATGAAA	AACCCAAATA	CGCTATCACT	ATTTTAG	237

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 159 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...159

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26

TTGGGTTTGG	TGWRCGGRAT	TTCTCTCTTG	CATTTGAGTT	TGGAGCAAAA	AATCAGCGTG	60
TTTCTTGGRC	YCAATTTAAT	GCTTTAYCCG	GTCAYAGAGG	TGCTTTTtag	TATCCTTAGG	120
CGCAAAATAA	AACGCCAGAA	AGCCACCCAT	GCCGGATAA			159

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1134 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...1134

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27

TTGGCTCAAC	CGGTTGAGGT	AAGAACAGTG	TTTATGTCCA	TGACTCTAAA	CGCTATGGGG	60
CAATTTGCCT	ATAATTTTCC	TGCTAATATC	AGCAAAGACA	AGCAAAAGCT	CACTATGGTT	120
TATATGGATA	AAGATTATGG	CGCTTATGGG	AATATTGTTG	CAATGGGTGG	GGAGTATGTC	180
AAGATTGAGC	TAGGGACAGA	TACAGGATTA	AATCCTTTTG	CTTGGGCAGC	TTGTGTGCAA	240
AAAACAAATG	CAACAATGGA	GCAAAAACAA	ACAGCTATTT	CTGTTGTCAA	AGAGCTTGTG	300
AAAAACTTAG	CAACTAAAAG	CGATGAAAAA	GATGAAAATG	GCAACAGCAT	CTCTTTTAGC	360
CTAGCAGATT	CTAATACGCT	TGCAGCGGCA	GTAACCAACC	TTATCACAGG	AGATATGAAC	420
CTAGATTATC	CTATCACTCA	ACTTATTAAT	GCTTTCGGGA	AAGACCACAA	TGATCCTAAT	480
GGGCTTGTCG	CGCGATTAGC	GCCTTTTTTG	AAATCAACCA	ATGGTGAATT	TCAATGGCTT	540
TTTGACAATA	AAGCAACAGA	TCGCTTAGAT	TTTTCAAAAA	CGATTATTGG	CGTTGATGGG	600
TCAAGTTTCT	TAGACAATAA	TGACGTTTCG	CCTTTTATTT	GTTTTTACCT	TTTCGCTCGT	660
ATCCAAGAAG	CAATGGATGG	GCGTAGATTT	GTCTTAGATA	TTGATGAAGC	GTGGAAATAT	720
TTAGGCGATC	CAAAGGTCGC	TTATTTTGTG	AGAGACATGC	TAAAAACTGC	AAGGAAAAGA	780
AACGCTATTG	TTAGACTTGC	GACTCAAAGC	ATCACTGATC	TTTTGGCTTG	CCCTATTGCT	840
GATACGATTA	GAGAACAATG	CCCTACAAAG	ATTTTTTTGA	GAAACGATGG	GGGTAATCTT	900
TCTGATTACC	AAAGATTAGC	CAATGTTACA	GAAAAAGAAT	TTGAAATCAT	CACCTAAGGGG	960
CTGGATAGGA	AAATCCTCTA	CAAACAGGAT	GGAAGCCCTA	GCGTTATCGC	TAGTTTTAAT	1020
TTGAGAGGCA	TTCTTAAAGA	ATATTTGAAA	ATTTTATCCA	CAGATACTGT	ATTTGTCAAA	1080
GAAATTGACA	AGATTATCCA	AAACCATAGT	ATCATAGATA	AATATCAGCC	TTGA	1134

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 465 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

89

- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...465
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28

ATGAAATTGG	TGAGTCTTAT	TGTAGCGTTA	GTTTTTTGTT	GTTTTTTAGG	GGCTGTAGAG	60
TTGCCTGGAG	TTTATCAAAC	TCAAGAATTT	TTATACATGA	AAAGCTCTTT	TGTGGAGTTT	120
TTTGAGCATA	ACGGGAAGTT	CTATGCCTAT	GGTATTTCTG	ATGTGRATGR	CTCTAAAGCC	180
AAAAAAGACA	AACTCAATCC	TAACCCAAAG	CTAAGGAATC	GCAGCGATAA	AGGCGTGGTG	240
TTTTTAAGCG	ATTTGATTAA	GGTTGGGGAA	CAATCTTATA	AAGGCGGTAA	GGCGTRTAAT	300
TTTTRTGACG	GCAAGACCTM	CCATGTGAGA	GTCACCTCAA	RTTCAAACGG	GGATTTGRAA	360
TTCACTTCAA	GCTATGRCAA	ATGGGGGTAT	GTGGGCAAAA	CCTTCACCTG	GAAACGCCTG	420
AGCGATGAAG	AAATCAAAAA	TCTAAAGCTC	AAGCGTTTTA	ACTGA		465

- (2) INFORMATION FOR SEQ ID NO:29:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 24 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: double
    - (D) TOPOLOGY: circular
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (iii) HYPOTHETICAL: NO
  - (iv) ANTI-SENSE: NO
  - (vi) ORIGINAL SOURCE:
    - (A) ORGANISM: *Helicobacter pylori*
  - (ix) FEATURE:
    - (A) NAME/KEY: misc\_feature
    - (B) LOCATION 1...24
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29

TTGGAGACCC TATTCTTGGT ATAG

24

- (2) INFORMATION FOR SEQ ID NO:30:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 345 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: double
    - (D) TOPOLOGY: circular
  - (ii) MOLECULE TYPE: DNA (genomic)
  - (iii) HYPOTHETICAL: NO
  - (iv) ANTI-SENSE: NO
  - (vi) ORIGINAL SOURCE:
    - (A) ORGANISM: *Helicobacter pylori*
  - (ix) FEATURE:
    - (A) NAME/KEY: misc\_feature
    - (B) LOCATION 1...345
  - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30

ATGSACACAC ACGACAGGCG CAAGTTAAGA ATTARCCTTA CACAAACGAC GACTTTAGTG

60

90

GCCACTATTG	GCTCAAACGC	CCCTTATATC	GGTCTTTTAG	GGACGGTTAT	GGGGATCATG	120
CTCACCTTTA	TGGATTTAGG	CTCAGCTTCT	GGCATTGACA	CTAAGGCGAT	CATGACTAAT	180
TTAGCCCTTG	CTTTAAAAGC	GACCGGCATG	GGGTATTGG	TAGCGATCCC	TGCGATTGTG	240
ATTTATAACT	TGTTAGTGAG	AAAAAGCGAG	ATTTTAGTTA	CCAAATGGGA	TATTTTCCAC	300
CATCCGGTTG	ATACGCAATC	CCATGAGGTT	TATAGCAAAG	CCTAA		345

## (2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 204 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...204
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31

ATGCAGGATT	TAGACAATAA	CATGTCCTTA	GACACCGCTC	ACAACACGCT	GAGTTCTAAC	60
GGGAAAAACA	TCACCATTCG	CGGGGTGGTA	AAAGCCTTAC	AAAAAATTGG	CGTGAGCGCT	120
AAGGGGATGG	TTTCAATCTT	GCAAGCCCTA	AAAAAAGCG	GCGCGATTAG	CGCGAAATGG	180
AGATACTATG	ATAAACAACA	ATAA				204

## (2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 267 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...267
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32

TTGCACCCTT	TAGCGGATGT	CTTTGTGGTG	AATGACAAAC	GGYCCGTTTT	AGCGATGGTA	60
RCGATGTTGA	TTGRCTCGTT	AGCGAATATC	TTTTTCAATT	ACTTGTTTAT	TTTTGKGTG	120
GAAGTGGGGG	TTCAAGGCAG	MGCGATAGTC	ACCGTGATAG	GGCATGCGAT	AGGGGGTTTA	180
GTCTTAATGC	AGCATTTTTG	GCGCAAAAAA	GGGGAGTTGT	ATTTTATCAA	ACTGATTTTC	240
TTTATCTTCA	GTCATTTCTT	CAGCTAA				267

## (2) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 831 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double



- (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...831
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33

ATGTTAAGGA	AAAAACATTT	AGCTTACTAT	GGGGCGAATT	TTCTCTTAAT	CATCGCTCAA	60
AGCTTGCCCC	ATGCGATTTT	AACCCCTTG	TTGCTTTCTA	AAGGGCTTAG	TTTGAGTGAA	120
ATCTTGCTCG	TGCAACCTT	TTTtagTTTT	TGCGTGCTGG	TGGCTGAATA	CCCAAGCGGC	180
GTTTTAGCGG	ATTTGATGAG	CCGGAAGAAT	TTATTCCTGG	TTTCTAATGT	GTTTTTAATC	240
GCTAGTTTTT	CGTTGGTGCT	GTTTTTTGAT	AGTTTTATCC	TCATGCTTTT	AGCGTGGGGG	300
TTGTATGGTT	TGTATAGCGC	ATGCTCTAGC	GGCACGATTG	AAGCTTCACT	CATCACAGAC	360
ATTAAGGAAA	ACAAAAAAGA	TTTATCCAAG	TTTTTAGCCA	AAAACAATCA	AATTACTTAT	420
TTGGGCATGA	TTATAGGGAG	TTCTTTGGGA	TCGTTTTTGT	ATCTCAAAGT	CCATGCGATG	480
CTGTATGTCG	TGGGGATTTT	TTAATCATG	CTCTGTGCGC	TAACAATCAT	CATTTATTTT	540
AAAGAAAAAG	AAGGGGATTT	TAAAAGCCAA	AAAAATTGA	AACTCCTTAA	AGAGCAAGTC	600
AAAGGCAGTC	TTAAAGAGCT	TAAAGATAAC	CCCAAGCTTA	AAATTTTGTT	AGTGGGGCAT	660
TTGATTACGC	CTGTCTTTTT	TATGAGCCAT	TTCCAAATGT	GGCAAGCGTA	TTTTTTAAAA	720
CAAGGCGTTA	AAGAGCAATA	CCTTTTTGTG	TTCTATATCG	CTTTTCAAGT	GATTTCCATC	780
CCTCATTCAT	TTTTTAAAAG	CCAAAAATTA	KCAGCCAAAA	AATCGCCCTG	A	831

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 282 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...282
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34

TTGTATCCGC	CGGGATCTGT	GGTTAAATG	GGCGTGGGGT	TAAGCTTTTT	AGAAAACCTT	60
CATATCACAG	AAAACACCAC	TATCCCCACA	CCGCCTTTTA	TTGAAGTGGG	CAAGCGCAAA	120
TTCAGGGACT	GGAAAAAAAC	AGGGCATGGC	AATTCTAATT	TGTATAAAGC	CATTAGGGAG	180
TCCGTGGATG	TGTATTTTTA	TAAGTTTGGG	CTTGAAATCT	CTATAGAAAA	MCTCTCTAAA	240
MCCTTTAAGG	RAAGTGGRCT	TTGGGGAAAA	AACGGRCGTT	GA		282

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 183 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

92

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...183
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35

ATGGCACACC	ATKAAGAACA	ACACGGCGGG	CACCACCACC	AYCACCACCA	CACACACCAC	60
CACCATTATC	ATGGCGGCGA	ACACCACCAT	CACCACCACA	GYTCTCATCA	TGAAGAAGGT	120
TGTTGCAGCA	CTAGCGATAG	TCATCATCAA	GAAGAAGGTT	GTTGYCACGG	GYAYCACGAG	180
TAA						183

## (2) INFORMATION FOR SEQ ID NO:36:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 894 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...894
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36

TTGGTTAAAA	TAAGGTTATT	TGATTTTACT	ATAAGGTTGT	TTAAACCTGA	ATTTACATT	60
TTTGATTTTT	TAAAAGGGAT	TAGAGTTCTT	ATGATTGAAT	GGATGCAAAA	TCATAGAAAG	120
TATTTAGTGG	TTACGATATG	GATAAGCACG	ATCGCTTTTA	TTGCCGCCGG	AATGATAGGT	180
TGGGGGCAAT	ACAGCTTTTC	TTTAGATAGC	GATAGCGCTG	CCAAAGTGGG	ACAGATTAAG	240
ATTTCTCAAG	AAGAATTAGC	CCAAGAATAC	CGCCGCCTTA	AAGACGCCTA	TGCTGAGTCT	300
ATCCCTGATT	TTAAAGAACT	CACCGAAGAT	CAAATCAAAG	CCATGCATTT	AGAAAAAAGC	360
GCGCTAGATT	CGCTCATCAA	TCAAGCTTTA	TTGAGGAATT	TCGCTTTAGA	TTTAGGGCTT	420
GGTGCTACCA	AGCAAGAAGT	GGCCAAAGAG	ATCAGAAAAA	CGAACGTTTT	TCAAAAAGAT	480
GGCGTTTTTG	ATGAAGAATT	GTATAAAAAT	ATCTTAAAC	AAAGCCATTA	CCGCCCCAAG	540
CATTTTGAAG	AAAGCGTTGA	AAGGCTTTTA	ATCCTTCAAA	AAATCAGCGC	TCTATTCCCC	600
AAAACCACCA	CCCCTTTGGA	GCAATCCAGT	CTATCGCTTT	GGGCAAAATT	GCAAGACAAA	660
TTAGACATTC	TTATCCTAAA	TCCTAATGAT	GTAAAAATCT	CTCTCAATGA	AGAAGAGATG	720
AAAAAATATT	ATGAAAACCA	TAGAAAGGAT	TTTAAAAAGC	CCACAAGCTT	TAAAACACGC	780
TCTTTATATT	TTGACGCTAG	TTTAGAAAAA	ACTGATTTGA	AAGAGTTGGA	GGAATACTAC	840
CATAAAAACA	AGGTGTCTTA	TTTGGACAAM	AGMGGGGAAA	TTACAGGATT	TTAA	894

## (2) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 273 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)

- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...273
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37

ATGGTTAAAC	ACTATCTTTT	CATGGCGGTT	TCGCAGGTCT	TTTTCTCCTT	CTTTT TAGTG	60
CTGTTTTTTT	TCTCTTCCAT	TGTGTTATTA	ATCAGTATTG	CAAGCGTAAC	GCTCGTGATT	120
AAAGTGAGCT	TTTTGGATCT	GGTGCAACTC	TTTTTGTATT	CCTTGCCWGG	AACCATTTTT	180
TTTATTTTGC	CGATYACTTT	TTTTGCGGCT	TKGCGYTTGG	GGSTTTCAAG	GCTTAGCTAT	240
GACCATGAAT	TGTTAGTGTT	TTTTCTCTCT	TAG			273

(2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 261 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...261
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38

ATGAGTAAAA	GAGCGATCCG	TTTCCCTAAC	AAGCTTTTTT	CATACCCTAA	ACCCAAAATA	60
AAAGCGACAA	ACACAAGCCA	CACCGTTTTA	TTGCGATACC	CGCTCAAACC	CCACGAAATG	120
GCCTTATTAG	CGCTCGCTAC	CTCACTGCTC	GCTCCAATTT	TTAACGCTAT	ACACAGCACT	180
AACGCGCTCA	ACGCTATCAA	ACCTGATGGC	ACCGGCTCTA	AAATTAACCC	TATAATCATG	240
CCCATGAAAA	TACAAAAATA	A				261

(2) INFORMATION FOR SEQ ID NO:39:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 426 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...426

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39

GTGTATAGCC	GTTTTTTTGC	CAACCAGCAT	GAATTTGACT	TTGAAGCTCA	AGGGGCGCTA	60
GGGAGCGATC	AATCAAGCTT	GAATTTCAAA	AGTACTCTAT	TACAAGATTT	GAATCAAAGC	120
TATAATTACT	TAGCCTATAG	CGCCACAGCA	AGAGCGAGTT	ATGGTTATGA	CTTCGCGTTT	180
TTTAGGAACG	CTTTAGTGTT	AAAACCAAGC	GTGGGCGTGA	GCTATAACCA	TTTAGGTTCA	240
ACCAACTTTA	AAAGCAATAG	CCAATCACAA	GTGGCTTTAA	AAAATGGCGC	GAGCAGTCAG	300
CATTTATTCA	ACGCTAACGC	AACGTGGAAG	CGCGTTATTA	TTATGGGGAC	ACTTCATACT	360
TTTATTTGCA	TGTGGGAGTT	TTACAAGAGT	TCGCTCACTT	TGGATCGAAT	GATGTGGCGT	420
CTTTAA						426

## (2) INFORMATION FOR SEQ ID NO:40:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 558 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...558

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40

ATGYATGAAA	ATGGTAGGGG	TGTACCTAAA	GATTACAAGA	AAGCGGTTGA	ATATTTCCAA	60
AAAGCTGTTG	ATAACGATAT	ACCTAGAGGG	TATAACAATT	TGGGCGTGAT	GTATAAAGAG	120
GGTAAGGGAG	TTCCTAAAGA	TGAAAAGAAA	GCGGTGGAAT	ATTTTAGAAT	AGCTACAGAG	180
AAAGGTTATA	CTAACGCTTA	TATCAACTTA	GGCATCATGT	ATATGGAGGG	CAGGGGAGTT	240
CCAAGTAACT	ATGCGAAAGC	GACAGAATGT	TTTAGAAAAG	CGATGCATAA	GGGCAATGTG	300
RAAGCTTATA	TTCTCCTAGG	GGATATTTAT	TATAGCGGAA	TGATCAATTG	GGTATTGAGC	360
CGGACAAAGA	TAAGGCTGGT	CCATTATAAA	ATGGCGGCCG	ATGTRAGTTC	TTCYAGAGCY	420
TATRAAGGGT	TGTCAGAGTC	YTATCSGTAT	GGGYTAGGCG	TGGAAAAAGA	KWAAAAAAG	480
GCGAAGAAT	ACATGCAAAA	AGCATGCGAT	TTTGACATTG	ATAAAAAATTG	TAAGAAAAAG	540
AACACTTCAA	GCCGATAA					558

## (2) INFORMATION FOR SEQ ID NO:41:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 420 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...420

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41

TTGCTCAACA	TGTGGGATGA	AGCCAAGAAA	GAAGGGATCA	ACATCAATAC	AGAAAAGCTC	60
------------	------------	------------	------------	------------	------------	----

TCTCAAGAAT	TGGGGGTTGT	GTGCGTGCCA	ACAAGCGCGA	GATYCAAAGA	AGATCGCTTG	120
AACACAGAGC	TTTTATTAGA	CGAAATTGTC	AGGCTTTATT	CTCAAAACAC	TACAAACAAC	180
GAAAACATCA	AAGTCCCATC	TCAAAGTTTT	AAAGAGTCTT	TAAAATACAG	CCAGAGCGCC	240
CAAAGAATCG	CTAAATCAGT	GATCAGTGAA	AACAAACAAA	ATGCGAGTTT	TGAACACACT	300
TATAAGATTG	ATAAGATTTT	TAATGCACCA	GCGTTATGGG	ATTTTCATTT	TTTTDGGGTT	360
TATGTTTATC	ATCTTTTCTT	TGAGCTTTTT	AATAGGAGGG	GGAGTGCRAA	AAGCCCTTGA	420

## (2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 582 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...582
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42

ATGCAAGAAG	CGTTGTTGCG	TTTTCAAGAG	GGCTTTAAGG	AGTGGGGTTA	TCTTATTTTA	60
TTTTTGTATT	CTTTGGGGGG	TGGGTATGTA	GGGATTGTCA	TCGCTTCCAT	TTTGAGCGCT	120
ACCACGCACG	CTTTGGATAT	AAAAATAACC	ATTCCTGTCTG	CTTTTTTAGG	GAATTTAATA	180
GGGAGTGGGG	CTCTTGTAAT	CTTTGCCCGC	TATCAAAAAA	GAGAGTTTTT	AAAGTATTTTC	240
CAAAAGCATA	GAAGAAAGCT	TGCTTTGGCG	AGTTTGTGGG	TGAAACGCTA	CGCCTTGCTC	300
ATGATTTTTG	TCAATAAATA	TCTCTATGGG	ATTAAAAGCG	TTGTGCCTTT	GGCAATTGGT	360
TTTAGCAAAT	ACCCTTTAAA	AAAGTTTTTA	TGGCTTAATG	TTTTTTCCAG	TTTTTTGTGG	420
GCGTTAATCG	TGGGGAGCGT	TTCTTTTCAA	GCGAGCGATT	GGGTGAAAAC	GCTGTATGAA	480
AGGCTTTCTC	ATTACACTTC	GTTTTTTGTC	ATAAGTTTTG	TTCTTATAGC	GCTTTTAATA	540
TGGTTTTTAT	TGAAACGATA	TTCGCGCAAA	ATGGGKTTTT	AA		582

## (2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 390 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...390
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43

ATGCGAAAGG	GGCGTGTGAT	GTTATGCGTG	TTTGATATAG	AAACCATTCC	TAATATAAGC	60
TTGTGTAAAG	AGCATTTTCA	ATTAAAAGAA	GACGATGCGC	TAAAAATCTG	TGAATGGAGT	120
TTTGAAAAGC	AAAAAGAAAA	AAGCGGGAGC	GAGTTTTTGC	CCCTTTATTT	GCATGAAATC	180
ATCTCTATTG	CAGCMGTCAT	TGGCGATGAT	TACGGGCAAT	TTATCAAAGT	AGGGAATTTT	240
GGTCAAAAAC	ACGAGAATAA	AGAGGATTTT	GCGAGCGAAA	AAGAGCTTTT	AGAGGACTTT	300

TTCAAATACT TTAACGAAAA GCAACCGCGC CTAATAAGCT TTAAWGCCAG GGGTTTTGGA	360
TATTCCCCTA CTCACGCTCA AAGCCCTTAA	390

## (2) INFORMATION FOR SEQ ID NO:44:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 924 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...924

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44

ATGGCTAAAA	AGAAAAATTGC	GATCAGCTGT	GGGGATATTC	AAGGCGTAGG	CTTAGAATTG	60
ATCTTAAAAA	GCCATAAGGA	AGTGAGTGCA	CTTTGTGAGC	CGTTGTATCT	CGTTCATAGC	120
GAACCTCTAG	AACGAGCCAA	TCAATTGCTT	GATAACGCTT	ATGAAACTAA	AACGCTTAAT	180
GCGATCGCTA	TTGATGCCCC	TTTACCCCTA	TTAAACTCTA	GCACGATAGG	CAAAGTCAGC	240
ACTCAAAGCG	GGGCGTATAG	CTTTGAGAGT	TTTAAAAAGG	CTTGCGAGTT	GGCGGATAGT	300
AAAGAAGTGG	ATGGCATTTG	CACTTTGCCT	ATCAACAAAC	TCGCATGGCA	ACAAGCTCAA	360
ATCCCTTTTG	TGGGGCATAAC	CGATTTTTTG	AAACAACGCT	ACAAAGATCA	TCAAATTATT	420
ATGATGCTTG	GGTGTTCAAA	ACTCTTTGTG	GGGCTATTTA	GCGACCATGT	GCCTTTAAGC	480
GCGGTTTCTC	AACTCATTCA	AGTGAAAGCG	TTAGTTAAGT	TTTTATTAGC	GTTTCAAAAA	540
AGCACTCAAG	CCAAAATCGT	TCAAGTGTGT	GGTTTCAACC	CCCATGCGGG	CGAAGAGGGA	600
TTGTTTGGGG	AAGAAGATGA	AAAGATTTTA	AAAGCCATTC	AAGAGAGCAA	CCAAACGCTA	660
GGTTTTGAAT	GCTTTTGGG	GCCACTGCCC	GCTGATAGCG	CTTTTGCCCC	CAATAAACGC	720
AAAATAACCC	CCTTTTATGT	GAGCATGAGC	CATGATGTAG	GGCTAGCCCC	TTTAAAGCG	780
CTCTATTTTG	ATGAAAGCAT	CAATGTGAGT	TTGAACGCTC	CCATTTTACG	CGCTTCCACT	840
GACCACGGCA	CGGCGTTTGA	TATTGCTTAT	CAAATAAGG	CGAACCATAA	AAGCTATTTG	900
AACGCGATCA	AATACTTGGC	TTAA				924

## (2) INFORMATION FOR SEQ ID NO:45:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 440 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...440

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45

ATGAGCAGCG	GGTTAATTTA	CATTTTCRTTA	GAAGTCTTGG	TARCGTGTTT	GATCACCGCT	60
CTAATCATGT	ATTATGTGAT	GAAAAAGATC	TATTACGCTA	GAGGGCAAGC	CATTTTAAAA	120
GGCGCTTCAG	CCAAAGCTAA	ATTAATGGAA	TTTCAAGCGA	AATCTTTCGT	GGAAGCTGAA	180

97

GAAATGCGCA	TGAAAAGCCA	AGAATGCAAG	TTGCAACAGC	AATATGAAAA	TAAGAATTTG	240
CAACTCCAAA	CCCATTTTGA	TAAAAAAGAA	GCGCATTTGA	AGCATTTAGA	AGCGCAGCAC	300
AAAGAATTTG	TAAGAGATGA	AAAACGCTAT	TTGGAAAAGG	AAAAAAAAGA	GCTTGAAAAA	360
GAACGCCAAA	TTTTAGAAMM	AGAGAGGGAA	AATTTTTRRR	RSCAGCGCGC	CTTTGTRRRG	420
RRGSCTCRRG	CCAAAGCGCT					440

## (2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 384 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...384
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46

ATGAATATCA	AAATTTTAAA	AATATTAGTT	GGAGGGTTAT	TTTTTTTGAG	CTTGAACGCC	60
CATTTATGGG	GGAAACAAGA	CAATAGCTTT	TTAGGGATTG	GTGAAAGAGC	CTATAAAAGC	120
GGGAATTATT	CTAAAGCGGC	GTCTTATTTT	AAAAAAGCAT	GCAACGATGG	GGTGAGTGAA	180
GGCTGCACGC	AATTAGGAAT	CATTTATGAA	AACGGGCAAG	GCACTAGAAT	AGATTATAAA	240
AAAGCCCTAG	AATATTATAA	AACCGCATGC	CAGGCTGATG	ATAGGGAAGG	GTGTTTTGGC	300
TTAGGGGGGC	TTTATGATGA	GGGTTTAGGC	ACGGCTCAAA	ATTATCAAGA	AGCCMTTGAC	360
GCTTACGCAA	GGCATGCGTT	TTAA				384

## (2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 351 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...351
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47

ATGGCGATAG	CCATTAAGGA	TTTATTGAGC	GCTTATAAAG	TCGTTTTACC	TTTGGATAAA	60
ATCAGCATGC	CATCTAGCGC	GGATTTGAAG	CTCACTTTGC	AATTCTTAAA	AAACACCGCC	120
CCCTTATTTA	GCGTTCAAGG	CAGCGTTAAT	TTGCAAGAAG	GCACTTTCTC	GCTCTATAAT	180
ATCCCCCTTT	ACACGCAAAG	CGCTCAAATC	AATTTGGACA	TCGCCCAAGA	ATACCAATAC	240
ATCTACATAG	ACACGATCCA	CACGCGCTAT	GCAAACATGC	KGGATTTAGA	CGCTAAAATC	300
GCTTTAGATT	TAGGTCAAAA	AAACCTYTCY	YKGGAKKCY	TAGGKCCATA	A	351

## (2) INFORMATION FOR SEQ ID NO:48:

98

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 249 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...249
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48

ATGCCGGATA	AYTTGCATTT	GCACACCCTT	TTATYTAAAT	TCTTGCAACA	ACGCTCTTTC	60
AATTACCCTA	ACCCTTTATG	CGCGTTTATC	CTTATTCTAT	GCAACCTGCC	TTTTATTTTA	120
ATAAGCGTTT	TGTTTCGCTT	GGACGCTTAT	GCGCTCATTG	TGATTAGCCT	AGTCTTTATC	180
RCATGCTATT	TAATAGGCTM	TGCTTATTTG	AATAGGCAAG	TTTGCCTTTT	AGAAAAGCGG	240
GCCTTTTAA						249

## (2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 351 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...351
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49

TTGAAAGTAA	CRAATCCGCA	TTTGT'TGGTG	GTAATCCAAG	ATTTAAACGC	TCGCATCGCT	60
TTAATGAAAC	TCTTATTCCA	AAACGTTAAG	AGCGCGAACA	AAGAATTGGT	TTTTTGCAAT	120
AAAGAAAAAC	GCTTGATAAG	GTCTTTTGAT	GCACAAAAAG	AATACGGCAT	CACGCCTGTA	180
GAAAATATTT	TAAGCGTTTT	AGACACCGCT	ATGAATCCTA	ACAGCGCGCT	TGTGATAGAC	240
AATCTCAACG	AAGCGAAAGA	ATTGCACGAC	AAAGTAGGGG	CGGAAAAGTT	AAAATCGTTT	300
TTAGAAAAAG	CCCMTAGACA	ACGAGCAGTA	TTGCGTCATT	TTTGCATG	A	351

## (2) INFORMATION FOR SEQ ID NO:50:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 597 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO



(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...597

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50

ATGAAAGAAT	CTTTTTACAT	AGAGGGAATG	ACTTGCACGG	CGTGTCTAG	CGGGATTGAA	60
CGCTCTTTAG	GACGTAAAAG	TTTTGTGAAA	AAAATAGAAG	TGAGCCTTTT	AAATAAGAGC	120
GCTAACATTG	AATTTAACGA	AAATGAAACC	AATTTAGACG	AAATTTTTTA	ACTCATTGAA	180
AAACTGGGTT	ATAGCCCTAA	AAAAACTCTA	GCAGAAGAAA	AAAAAGAATT	TTTTAGCCCT	240
AATGTAAAT	TAGCGTTGGC	GGTTATTTTC	ACGCTTTTTG	TGGTGTATCT	TTCTATGGGG	300
GCGATGCTTA	GCCCTAGCCT	TTTACCTGAA	AGCTTGCTTA	CGATTAACCA	TCATAGTAAT	360
TTTTTAAACG	CTTGCTTACA	GCTTATAGGC	GCGCTCATTG	TCAATGCATT	GGGGAGGGAT	420
TTTTACATTG	AAGGGTTTAA	AGCCTTATGG	CACAGACAAC	CCAACATGAG	CAGCCTTATC	480
GCCATAGGCA	CAAGTGCTGC	CTTAATTTCA	GCCTGTGGCA	ATTGTATTG	GTTTATACCA	540
ATCATTATAC	CGATCAGTGG	TCTTATGGGC	ATTATTATTT	TGAAAGCGTG	TGCGTGA	597

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 258 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...258

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51

GTGGGGATTG	TGCCGGATAA	TTTGTGGAAG	CTCAAACGCT	TCAATCAAGA	CTGGCGCGTT	60
GGGGACACGC	TCATTACTGC	TATTGGGCAA	GGCTCTTTTT	TAGCCACGCC	TTTGCAGGTG	120
TTAGCCTACA	CAGGACTCAT	TGCGACAGGC	AAACTGGCAA	CGCCTCATT	TGCTATCCAT	180
AACCAACAAC	CGCTCAAAGA	CCCCCTGAAT	AGGTTTTCAA	AAAAAGAAGC	TCCAAGCCTT	240
GCGCGTGGGC	ATGTATGA					258

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1032 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

100

(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...1032

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52

ATGCAAAACT	TATTGATACA	AGCAGAAAAT	GCAATCGCTC	TACTTTTTTT	GTAAATGAC	60
AAAAACCTAA	AAGGAAAAAT	AGATTTGATA	TATATTGACC	CTCCATTTGC	TACAAACAAT	120
CATTTTACTA	TCACAAATGG	TAGAGCAACC	ACAATTAGCA	ATTCTAAGAA	TGGCGATATT	180
GCTTATAGTG	ATAAAGTAGT	GGGTATGGAT	TTTATGGAAT	TTTTAAAACA	ACGCCTGGTA	240
TTGCTTAAAG	AATTGCTTTC	AGAACAAGGC	TCTATCTATG	TGCATACAGA	TTACAAGATA	300
GGGCATTATG	TCAAGGTAAT	GTTAGATGAA	ATATTTGGCA	TACAAAATTT	TAGAAATGAA	360
ATCACACGCA	TAAAGTGCAA	TCCTAAAAAT	TTTAAAAGGA	TAGGCTATGG	TAACATAAAA	420
GATATGATTT	TATTTTACTC	TAAAGGAAAA	AATCCCATTT	TTAACGAACC	TAAGATCCCT	480
TATACGCCAC	AAGATTTAGA	AAAACGATTC	CCTAAAATTG	ACAAAGATAA	AAGGCGTTAC	540
ACTACCGTTC	CAATACATGC	TCCAGGAGAA	GTGGAAAGTG	GCGAATGTTC	TAAAGCATTT	600
AAAGGTATGC	TACCTCCAAA	AGGGCGGCAT	TGGCGCACTG	ATATTGCCAC	ACTTGAGCGT	660
TGGGATAAAG	AAGGTTTGAT	TGAGTATTCT	AACAATAATA	ACCCTAGAAA	AAAAATTTAT	720
GCCTTAGAAC	AAGTTGGCAA	AAGACTCCAA	GACATCTCGC	AATTTAAACA	CCCACAATAT	780
CCAAGCTACC	CTACAGAAAA	AAACGCTCAA	TTATTAGACT	TAATCATTAA	AACCTCTTCT	840
AATAAAGATA	GTATTGTTTT	AGATTGTTTT	TGTGGTTCTG	GAACAACCTT	AAAATCTGCG	900
TTTTTATTGC	AACGAAAATT	TATAGGCATT	GATAATTCCG	ATTTGGCTAT	CCAAGCTTGC	960
AAAAACAAGC	TTGAAACAAT	AACAAAAGAC	TTGTTTGTTC	CTCAAAATTT	TTATGATTTT	1020
CTTGTTTTTT	AA					1032

## (2) INFORMATION FOR SEQ ID NO:53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 531 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...531

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53

ATGACAAGCG	TTGTCATCAA	GCCCCATGCC	TATGGCGAGC	AAGTCCAAGA	AATAGAAGAA	60
GAGTCAGATA	GCGATTATGA	AAAGAATAAC	GACCAGGAAG	CGATCAATTT	TGGTATCGCC	120
TTGCATAAGG	GATTGGAATA	CCAATACGCT	TACAACATTC	CTAAACAAAG	CGTTTTAGAA	180
TATTTAAACT	ACCACTATGG	TTTTTATGGT	TTGGATTACC	AAGCGTTAGA	AGAAAAGTTTA	240
GAGCTTTTTG	AAAACGATGC	AGGGATACAA	GCCCTTTTTTA	AAAATCATGC	CTTAAAGGGT	300
GAAGCGGCTT	TTTTATTCCA	AGGGGTGTGT	TCTAGGATTG	ATGTTTTATT	GTGGGATAGA	360
GGGCAAAATT	TGTATGTTTT	AGATTATAAA	AGCTCTCAAA	ATTACCAGCA	AAGCCATAAA	420
GCGCAAGTGT	CTCATTACGC	TGAGTTTTTG	CGAACTCAAG	SCCCCCATTT	TAAGATACAA	480
GCGGGCATT	TTTACGCTCA	TAAAAGACTG	CTTGAAAAAT	YATGGGKTTG	A	531

## (2) INFORMATION FOR SEQ ID NO:54:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 783 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...783
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54

ATGTCTGAGG	ATTTGCCTTT	TGCGAGCGAT	TCGCAATTCA	CTTACAATGG	GGTGAGCATC	60
ACGCGCCCCA	CTAATGAGGT	CAATGATGTG	ATCAGCGGGG	TTAATATCAC	TTTAGAGCAA	120
ACCACAGAGC	CTAATAAGCC	TGCCATTATC	AGCGTGAGCA	GAGACAATCA	AGCCATTATA	180
GACAGCCTTA	AAGAATTTGT	CAAAGCCTAT	AATGAGCTTA	TCCCTAAACT	AGACGAAGAC	240
ACGCGCTATG	ACGCTGACAC	TAAAATCGCC	GGGATTTTTA	ACGGCGTGGG	CGATATTCGT	300
GCCATTAGAT	CCTCTCTTAA	TAATGTGTTT	TCTTATAGCG	TGCATACGGA	TAATGGGGTA	360
GAAAGCTTGA	TGAAATACGG	GCTTAGTTTA	GACGATAAGG	GCGTGATGAG	TTTGGATGAA	420
GCTAAATTAT	CAAGTGCATT	AAATTCTAAC	CCTAAAGCGA	CTCAAGATTT	TTTCTATGGG	480
AGCGATAGCA	AGGATATGGG	GGGCAGAGAA	ATCCACCAAG	AGGGCATTTT	TTCTAAATTC	540
AATCAAGTCA	TCGCTAACCT	CATAGATGGA	GGGAACGCTA	AATTAAAGAT	TTATGAAGAT	600
TCCCTAGACA	GAGACGCTAA	AAGCCTGACC	AAAGACAAAG	AAAACGCTCA	AGAGCTTTTA	660
AAAACCCGCT	ATAACATCAT	GGCGGAGCGC	TTTGCCGCTT	ATGATAGTCA	AATCTCTAAA	720
GCCAATCAAA	AATTCAATTC	CGTGCAAATG	ATGATCGATC	AAGCAGCGGC	TAAAAAGAAT	780
TAA						783

## (2) INFORMATION FOR SEQ ID NO:55:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 438 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...438
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55

ATGCGCATCG	TATTTATGGG	AACGCCTAGT	TTTGCTGAAG	TGATCTTAAG	GGCGTTGGTT	60
GAAAATGAAG	ATAAAAAGAT	AGAAGTGGTG	GGGCTATTCA	CTCAAAGGGA	CAAACCTTTT	120
GGGCGCAAAA	AAGAATTGAA	AGCCCCAGAG	ACTAAAACAT	ACATTTTAGA	AAATCATTTA	180
AATATTCCCA	TTTTCCAGCC	GCAAAGTTTG	AAAGAGCCTG	AAGTTCAAAT	CTTAAAAGGT	240
TTGAAGCCTG	ATTTTATCGT	GGTGGTGGCT	TATGGTAAGA	TTTTGCCTAA	AGAGGTTTTA	300
ACCATCGCTC	CTTGCAATTAA	TTTGCATGCG	TCGTTATTGC	CCAAATACAG	GGGGGCTTCG	360
CCCATTTCATG	AGATGATACT	CAATGACGAT	AGGATTTATG	GCATAAGCAC	CATGCTTATG	420
GAKTTTGGA	TTGGATAG					438

## (2) INFORMATION FOR SEQ ID NO:56:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 747 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO

- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...747
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56

ATGCGTTTTT	ATTTTAAATT	CCTTTGGCTT	TTAGGGATT	TTCTTATTTT	TTATTTTTTA	60
GACATTAAAG	GCAGCTCTTC	TTATATCAGC	GACCGGGTTA	AAAGCGCCTT	GATGAGCGCG	120
AAAAACAGCT	TACTAGACAA	CGTTCAAGCG	TATTTTTTTC	AAGCCCAAAA	CATTAAGGAA	180
TTTCAAAAAG	AACGCTTGAT	TTTAGAAGCT	TTAAAACTAG	AAAACGCTGA	TTTGAAAGAG	240
CGTTTGAATA	GTATTTATCC	TTTAGAAAAT	CCAAAAATGA	CTTATACCCC	TACTTTTCATG	300
ACTTCATTCA	TCAATTTTGA	AGACACACAC	AGCGTTTCTC	TCAACCCCTAT	TGTGAATTTA	360
GAAGAAAACA	AGATTTATGG	CCTTGTCTCT	CACAACCAAG	CCATAGGCAT	TGCCCGTGCTA	420
GAAAAAGGGC	GCTTGAACGG	GTTTTTGAAC	GCCCACAAGC	GGTGTGCTTA	TAGCGTGATG	480
ATAGGCCAAA	ATCAAGTCTT	AGGCTTTATA	GGGACTAATT	TCAAGCAAGA	ATTAGTCGTG	540
GATTTTCATTG	TCCCAAGCGC	TGAAATCAAC	ATAGGCGATC	AAGTGCTAAC	GAGCGGGCTA	600
GATGGGATTT	TTGGAGCGGG	GGTGTGTGTG	GGTGAAGTTT	CAAGCGTTGA	AGATCATTAC	660
ACTTATAAAA	GCGCGGTGTT	GAAAAACGCT	TTTTTAAGCG	AAGCCAAACT	TTTAAGGCAT	720
GTGTTTTTAA	GCGGTGTGAA	AAACTAG				747

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 360 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...360
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57

TTGGCTTTAA	GATTGCCTTT	TTTGATCGCT	CACGTCATCA	ACATGTTTTT	ATTCTACCTC	60
ATAGGGCGAA	AGATTTTAAA	AAAGCCTAAA	GACGCTCTTT	ATGTGGTATT	GACTTACGCT	120
TTATTGCCTG	GGGTGAATCT	CTTTGCGATT	TTACTGGCTA	AAAGCGTGCT	GGTGTTAAGC	180
CTTGGGCTTT	TGATTAGCTA	TTTGTATATT	AAAACCCAAA	AAATCCCTTA	TTTAACCCCTT	240
AGCGCTTGCG	CGTTTTTAGA	CGGTGCGTTC	ATCCCGCTTT	TACTAGGGGT	TTTTGCCTAC	300
GCTTTAAGGA	AAACGGCTAT	TTTAAGAGCG	CGATCTTTGC	TTTGGTGTTT	TTAMTTGTGA	360

(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 327 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...327

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58

GTGAATTTAA	TGGACTATTT	TTCTAAAAGT	TTGTTTTTAA	ATTCATTGAA	CACGCAGCGA	60
TTGATCGTCT	CCAACAAATT	AGCGATTGAC	GTGCAATACG	GCATGCTCCA	AAGTGTCCGC	120
AAAAATTACC	CTGATGTGGT	GGATGGGGGT	GTTAGGGAGG	GGCCTTTTTG	GGTGTTAGCC	180
GGGGCYTAA	TGCCTTCAAT	TTTAATAGAA	ATTGGTTATA	ATTCCCATGC	GATAGAATCT	240
AAACGCATCC	AAAGCAAACC	GTATCAAAAA	ATCTTGGCTA	AGGGCATTGC	TGATGGCATT	300
GATAGTTTCT	TCAGCAAGAA	TGATTAG				327

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 474 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...474

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59

TTGGCGTCTC	GCTATTCTGT	GGCTGTTGGG	AATTTATTTT	CAGAGCATTT	GTATGATTTA	60
AGAAATGAAA	CCATGACCAA	TCTCATTGGT	TTTTTACTGG	TGTTGGCGTC	CATTTGGGTG	120
TTTTTTTGTAG	CTCTTGGAGT	GTTGCTAGGC	AAGATGTTAG	TCTTTAGCGG	TCTAGGCATT	180
ATAGACAAGG	CGTTAGGGTT	TATTTTTTCA	TGTTTGAAGA	CTTTTTTAGT	GCTTTCTTTT	240
ATCCTTTATG	CGCTCTCTAA	AATGGATTTA	ATGAAAGACG	CTAACGCCTA	TTTGCAAGAA	300
AAGAMCRCTA	TTTTCCCCAC	CATRAAAARC	RTCRCCTAGTA	AGATCATGCG	CCTTGATGGC	360
GTCAAACATG	TGGAGAAAAA	CCTTAAAGAC	AACCTTGAAG	AAATGAGCGA	TGAAGTTAAA	420
AATAAAGGAT	CTATTGATAA	CGCCAAAGAA	TCTTTTAATA	AGGGCTACGG	ATAA	474

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 246 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...246

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60

TTGAGCAAGC	AAAGCGCGGA	CATTGTGATC	ACTAATGACT	CCTTAAGCTC	TTTAGTCAAA	60
GTTTTAGCGA	TCGCTAAAAA	AACTAAAAGC	ATTACTTGGC	AAAATATCTT	GTTCGCTTTG	120
GGGATTAAGG	CGGTTTTTAT	CGTGCTAGGG	CTTATGGGGG	TAGCGAGCTT	GTGGGAAGCG	180
GTCTTTGGCG	ATGTGGGGGT	TACGCTTTTA	GCYTTAGCCA	ATTCCATKCG	CACGATGAGG	240
GCTTAA						246

## (2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 240 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...240

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61

ATGAAAAATT	TAAGGCATTT	TAGAAAGCTT	ATCGCCTTTT	TAGGTTTTTC	ACCTCTTTTA	60
TTACAAGCGG	ATATGACTAC	CTTTTTTAAT	TCCATTGAAC	AACAGCTCAC	TAGCCCTACG	120
GCTAAAGGCA	TTTAAATGGT	TATTTTTTTA	GGACTTGCTA	TTTTTATATG	GAAAACTTA	180
GATAGATGGA	AAGAAATTTT	AATGACCGTG	CTTGCTTTAA	AAGRAGTCCC	CATGCAATMW	240

## (2) INFORMATION FOR SEQ ID NO:62:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 978 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...978

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62

TTGGCGGGTT	TGCYAGTGGG	GTGTATRCGG	ATGAAACAAA	CATTTTGGGR	ACTTAGTTGG	60
GGGGAAAAAA	GCCAAAAGGT	ATGCGTGCAT	CGTCCATGGT	ATGCTATATG	GAGTTGCGAT	120
AAATGGGAGG	AAAAAACACA	ACAATTTACA	GGAAACCAAC	TCATCACAAA	AACTTGGGCA	180
GGGGGTAATG	CGGCTAACTA	CTACCACTCT	CAAAACAACC	AAGACATCAC	AGCCAATTTA	240
AAAAATGATA	ACGGCACTTA	TTTTTTAAGC	GGTCTGTATA	ACTACACCGG	AGGGGAATAT	300
AATGGGGGGA	ATTTAGACAT	TGAATTAGGC	AGTAACGCTA	CTTTTAATCT	AGGTGCGAGT	360
AGTGGGAATA	GCTTCACTTC	TTGGTATCCT	AATGGGCATA	CTGATGTTAC	TTTTAGCGCT	420
GGGACTATCA	ATGTGAATAA	CAGCGTAGAA	GTGGGCAATC	GTGTGGGATC	GGGAGCTGGC	480
ACGCACACCG	GCACAGCCAC	TTTAACTTG	AACGCTAATA	AGGTTACTAT	CAATTCCAAT	540

105

ATCAGCGCGT	ATAAACTTC	GCAAGTGAAT	GTAGGCAATG	CTAACAGCGT	TATTACCATT	600
AATTCGGTTT	CTTTAAATGG	GGAACTTGC	AGKTCTTTAG	CTAGGGTGGG	CGTAGGGGCT	660
AATTGCTCCA	CTTCTGGGCC	TAGCTATTCT	TTTAAAGGGA	CSACTAACGC	TACTAACACS	720
ACTTTTAGCA	AWTCAAGCGG	SAGTTTCACT	TTTGAAGARA	ACGCCACTTT	TAGCGGGGCG	780
AAATTAAATG	GGGGGGCATT	CACTTTCAAT	AAAAAGTTTA	ACGCTACCAA	TAATACCGCT	840
TTTAATAGCG	GTAGTTTTAC	TTTTAAAGGC	ACAAGCTCTT	TTAATGGTGC	GAATTTTAGT	900
AACGCTTCCT	ATACTTTTAA	TAATCAAGCC	ACTTTCCAAA	ACAGCTCCTT	TAATGGGGGG	960
ACTTTTACTT	TTAATTGA					978

## (2) INFORMATION FOR SEQ ID NO:63:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 816 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...816

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63

TTGTAAAGTT	TAGTTAAAGG	GAAAACCATG	CTCCGCTCTC	TCTATAGTGC	CACTTCAGGG	60
ATGCTCGCCC	AACAAACGCA	CATTGACACC	ACTTCAAACA	ACATCGCCAA	TGTCAATACC	120
ACCGGGTTTA	AAAAATCTCG	CGCGGATTTT	AACGACTTGT	TTTACCAAGC	GATGCAATAC	180
GCCGGCACCA	ACACAAGCAA	CACGACTTTA	TCGCCAGATG	GCATGGAAGT	GGGCCTTGGC	240
GTACGCCCTA	GTGCGATTAC	CAAAATGTTT	TCGCAAGGCA	GCCCTAAAGA	AACGGAGAAT	300
AATTTAGATA	TTGCTATTAC	AGGTAAAGGC	TTTTTTCAAG	TCCAGCTTCC	TGATGGCACT	360
ACCGCTTACA	CAAGGAGCGG	GAATTTCAAG	CTAGACGAGC	AGGGCAATCT	TGTAACAAGC	420
GAGGGCTATC	TCCTCATCCC	TCAAATCACT	TTACCCGAAG	ACACCACGCA	AGTGAATATC	480
GGTGTGGATG	GCACGGTGAG	CGTGA CTCAA	GGCTTGCAAA	CGACTTCTAA	CGTGATCGGG	540
CAAATCACTT	TGGCTAATTT	TGTCAATCCG	GCGGGGCTTC	ATTCTATGGG	GGATAATTTG	600
TTTTCCATCA	CCAACGCTAG	CGGCGATGCG	ATTGTGGGCA	ACCCGGATTG	TCAAGGCTTA	660
GGCAAGTTAA	GGCAAGGCTT	TTTGGAGCTT	AGTAACGTGA	GATTGGTAGA	AGAAATGACA	720
GATCTAATCA	CCGCTCAAAG	GGCTTATGAA	GCCAAATCTA	AAAGCATTCA	AACCGCTGAT	780
GCCATGCTCC	AAACAGTCAA	TTCCCTCAAA	CGCTAA			816

## (2) INFORMATION FOR SEQ ID NO:64:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 273 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...273

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64

ATGCAAAATG	GGTATTATGC	GGCCACAGGG	GCAATGGCTA	CACAATTTAA	CCGCTTGGAT	60
TTAACCTCTA	ACAATTTAGC	CAACCTAAAC	ACCAACGGCT	TTAAAAGAGA	CGATGCGATT	120
ACAGGCGATT	TTTTAAGGCT	TTACCAAGAA	TACCGAGAGC	AACTGCCCTT	AGAAGATCAA	180
ACCAAAGCGA	GCGCGAAGTA	TCTCAACCGC	AMCCTCAATC	GTGTGCCTAT	TCTATCARAA	240
ATCTATACKK	ATAGGRAGCT	TGGCYGCGTT	TGA			273

## (2) INFORMATION FOR SEQ ID NO:65:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 585 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...585
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65

GTGGGGGCTA	TGCCTACTAT	CCAAATCCGT	GRCTTTGGAG	CGGGGGGTTT	AGGGCATAGC	60
GATGCGACGC	TCATGTTAGT	TAATGGTATT	CCTGTTTATA	TGGCCCCTTA	CGCTCACATT	120
GAGCTAGACA	TTTTCCCTGT	TACCTTTCAA	GCCATTGATC	GCATTGATGT	GATCAAAGGT	180
GGAGGCAGCG	TGCAATATGG	GCCTAACACT	TATGGGGGTA	TTGTCAATAT	CATCACTAAA	240
CCTATCCCTA	ATCAATGGGA	AAACCAAGCG	GCTGAAAGGA	YCACTTATTG	GGCTAAGGCT	300
AGAAACGCTG	GGTTTGCCGC	TCCCCYTGAT	AAAACCGGGC	ATCCTTCTTT	CATCAAGTCT	360
TTAGGCAACA	ACCTCCTCTA	TAACACTTAT	GTGAGGAGCG	GAGGGATGAT	CAATAAGCAT	420
GTGGGTATCC	AGCGCAAGCT	AACTGGGTTA	GAGGCCAAGG	CTTTAGGGAC	AATAGCCCCT	480
CTAGTATTTT	AAACTATTGG	CTGGATGGGG	TCTATGACAT	CAATGAAAGC	AATGGGATTA	540
AAGCCTATTA	CCAATACTAC	GATTTTGGCT	ATCGSCCAAC	CGGGA		585

## (2) INFORMATION FOR SEQ ID NO:66:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 255 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...255
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66

ATGAGAWAGG	AGAAAATAAT	GACGAATTTT	GAAAAGRTTA	TCGCGCAAAA	CAGGCTCAAA	60
ACGAACGCGG	TTTTAACCAC	TTACTGCGCG	ATTTTTGCTT	TTATTGGGTT	GTTGGTGGAT	120
GCTATTAGAA	TCAACGCTAA	TGATTTAGGT	ATAGCCCTTT	TTAAACTCAT	GACTTTTCAA	180
ATTTTTCTTA	CGRTTACTAT	TGTCATGTTT	GTGGTGGCTT	TTGTCATTAK	KCKKAGTTTG	240
TATCCAAAAT	TTTAG					255



## (2) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 231 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...231
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67

ATGGRCATGT	CTCATATTAT	TAAGAGCATT	GAAGCTTTAG	ATGACTATAC	CATTAGATTC	60
ACGCTTAATG	GGCCAGAAGC	CCCGTTTTTA	GCGAATTTGG	GCATGGACTT	TTTAAGCATT	120
TTGAGTAAGG	ATTACGCTGA	TTACTTGGCT	CAAAATAATA	AAAAAGACGA	GTTGGCTAAA	180
AAMCCTGTTG	GGACAGGGCC	TTTCAAATTC	TTTTTGTGGA	ATAAAAGATG	A	231

## (2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 591 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...591
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68

TTGATGAGGA	AAATTTTTTC	TTATATTCT	AAGGTTCTAT	TATTTATTGG	GGTGGTTTAT	60
GCAGAGCCTG	ATTCTAAAGT	GGAAGCCTTA	GAAGGGAGGA	AGCAAGAGTC	TTCTTTGGAT	120
AAAAAAATCC	GCCAAGAATT	GAAGAGTAAG	GAATTGAAGA	ATAAGGAATT	AAAGAACAAG	180
GATTTGAAAA	ATAAAGAAGA	AAAGAAAGAA	ACAAAAGCCA	AGAGAAAACC	CAGAGCAGAA	240
GTCATCATG	GGGACGCCAA	AAATCCCCT	CCAAAGATCA	CGCCTCCTAA	AATCAAAGGG	300
AGTAGTAAGG	GCGTTCAAAA	TCAAGGCGTT	CAAAACAACG	CGCCAAAACC	TGAAGAAAAA	360
GATACAACCC	CTCAAGCTAC	TGAAAAAAT	AAGGAAACAA	GCCCTAGCTC	TCAATTCAAT	420
TCCATTTTTG	GTAATCCTAA	TAACGCTACC	AACAACACCC	TTGAAGATAA	GGTCGTAGGG	480
GGCATTTCAT	TGCTTGTTAA	TGGTTCGCCT	ATCACGCTGT	ATCAAATCCA	AGAAGAGCAA	540
GAAAAATCTA	AAGTGAGYAA	SGCTCRAGCT	AGGGATCGTT	TGAKTCKCTG	A	591

## (2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 540 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double

- (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...540
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69

ATGAAGAGAT	CTTCTGTATT	TAGTTTCTTG	GTAAGCTTTT	TATTGGTAGT	TGGCTGTAGT	60
CATAAAATGG	ATAATAAGAC	TGTGGCTGGC	GATGTGAGCA	CTAAAGCGGT	TCAGACTGCG	120
CCTGTTACTA	CAGAACCAGC	TCCAGAGAAA	GAAGAGCCTA	AACAAGAGCC	AGCTCCAGTG	180
GTTGAAGAAA	AGCCGGCTAT	TGAAAGCGGG	ACTATCATCG	CTTCTATTTA	TTTTGATTTT	240
GACAAGTATG	AGATCAAAGA	ATCCGATCAA	GAGACTTTAG	ATGAGATCGT	GCAAAAAGCT	300
AAAGAAAACC	ACATGCAAGT	GCTTTTGGAA	GGCAATACCG	ATGAATTTGG	CTCTAGCGAA	360
TACAACCAAG	CGCTTGGCGT	TAAAAGGACT	TTGAGCGTGA	AAAACGCTTT	AGTCATTAAA	420
GGGGTAGAAA	AAGATATGAT	CAAAACCATC	AGTTTGGCG	AAAGCAAACC	CAAATGCGTC	480
CAAAAACCTA	GAGAATGTTA	CAGAGAAAAC	AGAAGAGTGG	ATGTCAAATT	AGTGAAGTAA	540

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 861 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...861

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70

ATGGGAACGC	TCATTGAAAA	ATGGTTTGGC	TTCTCTCAAA	TCAGAGAAGA	ATTAGAAGCT	60
CGCATCAGTG	AGTTAGAAGA	CGAAAACACC	GAATTGTTAA	GAGAAAGAGA	ATACTTAGCT	120
GCAGAACTA	GCGAGTTAAA	AGACGCTAAC	GATCAATTAC	GGCAAAAAAA	CGACAAGTTA	180
TTCATAACAA	AAGACAAGCT	AACCAAAGAA	AACACCGAGT	TATTCGCAGA	AAACGAAAGC	240
TTATCTGTAA	AAATCAGCGG	GTTAGAACAC	TCTAACGATC	AATTATGGCA	AAACAATAAC	300
AAGCTCACTA	AAGAAAAAGC	AGAAGTGAAG	ACGGAAAAAG	ACATTTTAGC	TAAAGAAAAAC	360
ACACGCTTAT	TAGCAGCCAG	AGATCGGCTG	ACTGAAGAAA	AAAGAGAATT	GACAACAGAA	420
AAAGAAAAGC	TAAAAAGAGA	AAACACCGAG	CTAACCCATA	AAATCACCGA	GCTGACTAAA	480
GAAAAATAAG	CACTAACCAC	CGAAAACGAC	AAGCTCAACC	ACCAAGTTAC	CGCGCTCACT	540
AATGAGCGAG	ATAGTCTCGA	ACAAGAGCGA	GCGCGATTGC	AAGATGCGCA	TGGGTTTCTA	600
GAAAAACGAT	GCACCAATTT	AGAGAAAGAA	AACCAACGCC	TAACTGACAA	GCTCAAACAA	660
TTAGAAAGCG	CTCAAAAAAG	CTTGGAAGAA	ACTAACAATC	AATTACGGCA	AGCTTTTAGAA	720
AACTCTAATG	TCCAATTAGC	ACAAGCTAAA	GARARAATWG	CCATAGAGRA	AAGCGAGCTG	780
GMGCGAAGAA	ATCGCACGCT	TGAAGAGCTT	AGAGGGTATG	GAAGCCAAAA	GSCGATCTGG	840
ACTTACACAW	CAGGCGTTTA	G				861

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:

109

- (A) LENGTH: 333 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...333

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71

GTGTTGCGCA	AGCTTTTGGG	TAAAAATTGC	ATAGAAACGC	ATAAGGGGGT	GGGCTATCGC	60
TTAACCCACT	ATGAAAAAAAA	ATCCCTCAA	CTCTTTTGTAG	GGACTTATTT	AGGCTCTTCG	120
TTTGTGTAA	TGCTAGTGAT	TAGCGTTTTA	GCGTTTAACT	ATGAAAAAAAA	CGAAAAAATC	180
AAARTGATAC	GCATGGACAT	GGACAAAATG	GCTTCTAAGA	TCGCTAGTGA	AATTATCCAA	240
TTGCACATGC	AAACGCATGC	GGATTATCAC	AACGCTTTAA	ACGCCCTGAT	TTCACGTTAT	300
AAAGACGTTT	CCATAGYCCT	YTTYGAYACG	TAA			333

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 375 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72

TTGATGACCA	AAAGCTTAAA	ACTCATTCAA	AAAGGGGTTA	AAAACCTCTA	TGAAACCCTT	60
AAAAATAGGG	CTTTAGAGCA	TCAAGACACG	CTAATGGTGG	GCAGAAGCCA	TGGGGTGT TT	120
GGCGAACCCA	TCACCTTTGG	TTTAGTTTTA	GCTCTTTTGT	CTGATGAAAT	CAAACGGCAT	180
TTAAAAGCCC	TGGATTTAAC	GATGGAATTT	ATCRGCGTWG	GGGCGATCAG	TGGGGCTATG	240
GGGAATTTTCG	CGCACGCCCC	TTTAGAATTA	GAAGAATTAG	CGTGCGGATT	TTTAGGCTTA	300
AAAACCGCTA	ATATCAGCAA	TCAAGTCATT	CAAAGAGACC	GCTACGCAGG	CTTGCATGCG	360
ATCTGGCTCT	TTTAG					375

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...288

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73

TTGTCAGACG	CTTCAAAAAG	ATCCCTTAAT	CCAACCTTAA	TGATGAATAA	TAATAATACC	60
CTACCCAAAC	CCCTAGAAGA	AAGCCTAGAT	TTAAAAGAGT	TTATCGCTCT	TTTTAAAACC	120
TTTTTCGCAA	AAGAAAGAGG	TTCTATTGCT	TTAGAAAACG	ATCTCAAACA	GGCTTTCACT	180
TATTTAAATG	AAGTGGATGC	GATCGGTTTG	CCTGCCCCCM	AAAAGCGTGA	AAGAAAGCGA	240
TCTTATTGTT	GTCAAATCA	CCAAATTAGG	GACGCTCCAT	TTAGATGA		288

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 243 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...243

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74

TTGCCTATTA	TTTAYCTGT	AATCGTGATG	ATGTTTTTTT	CCAAAATCGT	TGGCGATTTT	60
ATTGAAAAGC	ATTATCGCGT	CAAAACTTTA	GCCTTTGTGT	TTTTGCTCGT	TGTGGGCGTG	120
TTTTTGTTTT	TAGAAGGCTT	GCATTTACAC	ATCAATAAAA	ACTATTTGTA	TGCGGGTATT	180
GGTTTTGCCT	TGCTCATAGA	ATGCTTGRAT	ATTTTCATAG	AAAAGAAAAT	GAAAAAAAGT	240
TAA						243

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 798 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...798

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75

111

ATGATAAAAG	CGCGGTTTAA	AAAACGCCTT	TTAGGATCTA	GGGGCGCGTT	TGATTTGAAT	60
ATAGACTTAG	AAATTAAAGA	AGCAGAAAGT	GTCGCTTTAT	TAGGAGAATC	GGGAGCGGGT	120
AAAAGCACGA	TCTTACGCAT	TTTAGCAGGG	CTTGAAGCGG	TGAGTAGCGG	CTATATTGAA	180
GCCAATCATT	CAGTATGGTT	AGACACTCAA	AAAAAGATTT	TTTTAAAACC	ACAACAGCGA	240
AAAAATCGGCT	TTGTGTTTCA	AGATTACGCC	CTATTCCCTC	ATTTAAACGT	GTATCAAAAC	300
ATCGCCTTTG	CTCACCTTAA	AGATAAAAAT	AAAATCCACG	AAGTGTTACG	CTTAATGCGT	360
TTAGAAAACC	TAAGCCAGCA	AAAAATTCCC	AAACTCTCTG	GCGGGCAAGC	CCAACGAGTC	420
GCTTTAGCAA	GAGCTTTAAT	CGCAGCCAAA	AATCTATTGC	TTTTAGATGA	GCCTTTAAAC	480
GCCCTAGATA	ACGCCTTAAA	AAACGAGGTG	CAACAAGGTT	TGCTTGATTT	TATCAAGCGT	540
GAAAATTTAA	GCGTGTTATT	GGTAAGTCAT	GATCCAAACG	AAATAACCAA	ACTCGCGCGA	600
ACTTTCCTCT	TTTTAAACAA	TGGCGTTATT	GATCCTAATC	AAGAAAATCG	GCTTTTTTCA	660
AACCGCTTAT	TGGTAAACC	TCTCTTTGAA	GATGAAAATT	ATTGCCATTA	TGAGGTCATT	720
CCTCAAACGA	TCAGTTTGCC	CAAAGATTGT	CTGAACCCAA	CTTTTAAGCT	TGATTTTCATT	780
CAAAACAAAA	AATTTTAG					798

## (2) INFORMATION FOR SEQ ID NO:76:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 195 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...195
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76

GTGAAATTCA	GCGTTTTAAC	CCTTTTCCCG	CAACTCATCT	TGCCTTATTT	TGAAGATTCT	60
ATTTTAAAAA	GAGCGTTAGA	AAAAAACCTT	TTTGAATTGG	AAGTGTTAAA	CCTTAGAGAT	120
TTTAGCGCTA	ACAAATATCA	AAAAGCGGAK	TCACACGCTC	ATTGGTGGGG	GTGCGGGGCA	180
AATTTTAGAC	CCTGA					195

## (2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 414 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...414
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:77

TTGTGGCGCA	CACCGAAGAC	RCCCTTAGTC	ATTAAACCCT	ATTTGAAAAG	CATGAGCGAT	60
TCAGAGATYT	TTGCGGYCAY	GTGCGTGGGY	ATGGCYAGCG	TTRCGGGGCC	TGTGTTAGCC	120

112

GGGTATGCGA	GCATGGGCAT	TCCTTTACCT	TATTTAATCG	CCGCATCGTT	TATGTCCGCT	180
CCTGGGGGGT	TGCTGTTTCG	TAAAACCAAT	TACCCGCAAA	ACGAAACCAT	TTCTAGCCAT	240
GCAGATGTTT	CTGCAGAAGA	GCATGTCAAT	ATTATAGAAG	CTAYCGCWMA	TGGGGCAAGC	300
ACAGGGGYTC	ATTTAGCCTT	GCATGTGGGG	GCGATGCTTT	TAGCCTTTGT	GGGGATGGTC	360
GCGCTCGTTA	ACGGGCTTTT	AGGGGTTGTA	GGGGGATTTT	TAGGCATGGA	GCAT	414

## (2) INFORMATION FOR SEQ ID NO:78:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 348 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...348

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78

GTGATGA	ACT	TTTTT	TG	TGGG	CGG	ACTTT	TCC	ATTG	TTT	TG	TGTA	ATGT	GGT	GGT	CAT	CACTT	TAC	60
TCCG	CGCT	TCC	ACC	TAC	AGC	CC	TGT	AG	AA	GGT	GC	AGA	AG	AT	ATT	GTT	CA	120
CATT	TG	ACCA	GTTT	CT	AT	GG	GCC	AG	CG	ACT	GGG	TT	ATT	GT	TTG	GKTTT	TAC	180
GCC	GCT	AT	CA	ACC	AC	ACTT	TG	GGT	TT	GG	AT	TGG	AG	AC	CT	ATT	CT	240
G	TAG	CG	AT	CA	AC	ACT	GT	TCC	TG	CT	GC	GA	TT	TAT	CCC	ACT	AT	300
C	AC	AA	AG	TG	T	TAG	GC	AT	C	TGA	AG	GC	AT	T	GC	AT	T	348

## (2) INFORMATION FOR SEQ ID NO:79:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 684 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...684

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79

GTG	CTTTT	TGG	GCA	AG	CAT	AG	TGG	GGG	CGGG	TTG	TTG	AG	CG	CTT	TAW	CGC	GTT	AAG	TTT	60
GG	AT	CT	TGGG	TGG	TG	AG	TAT	CCA	AG	CG	TTA	GAG	TG	CG	GAG	TA	ACT	TCT	AA	120
TT	AG	AAT	TGG	T	TTT	TG	TGA	AA	ATTT	CCCT	AAAA	AG	CT	CA	GCG	CGT	TCC	G	CT	180
GGG	TT	AG	AAA	AT	AT	CC	AAA	GG	ATTTT	TAAG	AAG	TGG	CT	TG	AAT	AG	CCCC	AT	G	240
GAT	G	CGG	CG	TTTT	T	AT	CA	TAA	AG	AG	TG	TT	AC	AG	CT	T	AG	AAAA	AG	300
AC	CC	CT	C	ACC	CT	AA	AG	AG	TT	TT	AT	CG	TT	A	TG	AA	AT	CA	G	360
CT	AG	A	TT	TAC	TAG	AC	A	TAA	ACT	AG	AA	TC	GA	AG	TT	T	CT	C	AAA	420
GT	GG	TTT	TG	TTTT	AA	AG	GG	GG	CT	A	AT	AC	CT	A	TC	G	CT	C		480
AA	CA	AT	T	TAG	GG	AG	CG	T	GG	C	AAA	GC	AG	CG	AG	T	G	T	AG	540
AT	T	G	T	A	AG	CC	T	ACT	TT	T	CT	C	A	CG	CT	T	T	A	G	600

113

GCGCACGCCC TAGCGGGTTT ARAATTTAAG AATMATTAMG CTTTAACGCC CSTAGATTTG 660  
 ATAGAAAAGR TCAAACGACT ATAA 684

## (2) INFORMATION FOR SEQ ID NO:80:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 328 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...328

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80

GTGCTTTTAT	ACCTAGCACT	AACCTTGAGT	TTAGGCATTG	CTATGCTTTT	AGTGGAATG	60
CTGATTGGAA	ATTTGGGTAA	AAAAGACGTT	GTTTCCAATT	ATCAAATCTT	AGATCCTAAA	120
AGGAAAAAAT	ATTACCCTTT	CACCTCTTTT	TTTATTTTAG	GCGGCCCTCT	CATTCTATCT	180
TTTTATGCGG	TGGTGTTAGG	CTGGGTGCTT	TACTATCTTT	TTGTAGTAAC	TTTGTATTG	240
CCTAAAGATT	TAGMGCAGGC	TAAAATGCAR	TTCMGMATGC	TTCAAAATGG	CAGTTTGATC	300
TGGCCTGTTA	TTGACTTTAG	CGCATGCT				328

## (2) INFORMATION FOR SEQ ID NO:81:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 294 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...294

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81

TTGACAACAA	AAGCGTGTTG	GTTGCTTCGG	GTTTGTGTTT	ATAGAAGTCT	AAATATTACA	60
ATCAAGGATA	GAACGATGAA	AACCAATGGT	CATTTTAAGG	ATTTTGCATG	GAAAAAATGC	120
TTTTTAGGCG	CGAGCGTGGT	GGCTTTATTA	GTGGGTGTA	GCCCGCATAT	TATTGAAACC	180
AATGAAGTTG	CTTTGAAATT	GAATTACCAT	CCAGCTAGCG	AGAAAGTTCA	AGCGTTAGAT	240
GAAAGATTTT	TACTTTTAAG	GCCAGCTTTC	CAATACAGCG	AKAATATTTG	CTAA	294

## (2) INFORMATION FOR SEQ ID NO:82:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 438 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

114

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...438
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82

TTGAGTGAGT	GGCAAACATT	TTGTTTAAAA	GATTTAGGGA	AAATAGTCGG	CGGCGCTACC	60
CCACCTACCA	ATAACCCCAA	AAATTATGGC	AATAAAATTC	CTTGGATTAC	CCCTAAAAGAT	120
TTATCCACTT	TACAAGGGCG	CTACATTAAA	AAAGGCAGCC	GCAGCATTTT	ACGATTAGGG	180
TTTAAATCAT	GCTCTTGTTG	GTTGCTCCCA	AAGCATGCCA	TTTTATTTTC	TTCAAGAGCT	240
CCCATAGGTT	ATGTGGCAAT	TGCTGAAAAA	AGGCTATGCA	CCAATCAAGG	TTTTAAAAGT	300
ATTATCCCTA	ACAAAAAAT	TTATTTTGAA	TTTTTATATT	ACTTATTAAA	ATACTATAAG	360
GATAACATTT	CCAACATAGG	GGGCGGAAC	ACTTTTAAAG	AAGTTTCAGG	GGCTACTTTA	420
GGKTCTATTC	CAAGTTAA					438

## (2) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 822 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...822
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83

ATGGAATTTA	TGAAAAAGTT	TGTAGCTTTA	GGGCTTCTAT	CCGCGGTTTT	AAGCTCTTCG	60
TTGTTAGCCG	AAGGTGATGG	TGTTTATATA	GGGACTAATT	ATCAGCTTGG	ACAAGCCCGT	120
TTGAATAGCA	ATATTTATAA	TACAGGGGAT	TGCACAGGGA	GTGTTGTAGG	TTGCCCCCA	180
GGTCTTACCG	CTAATAAGCA	TAATCCAGGA	GGCACCAATA	TCAATTGGCA	CTCCAAATAC	240
GCTAATGGGG	CTTTGAATGG	TTTTGGGTTG	AATGTGGGTT	ATAAGAAATT	CTTCCAATTC	300
AAGTCGCTAG	ATATGACAAG	CAAGTGTTTT	GGTTTTAGAG	TGTATGGGCT	TTTTGATTAC	360
GGGCATGCCG	ATTTAGGTAA	ACAAGTTTAT	GCACCTAATA	AAATCCAGTT	GGATATGGTC	420
TCTTGGGGTG	TGGGGAGCGA	TTTGTTAGCT	GATATTATTG	ATAAAGACAA	CGCTTCTTTT	480
GGTATTTTGT	GTGGGGTCGC	TATCGGCGGT	AACACTTGGG	AAAGCTCTGC	AGCAAACAT	540
TGGAAAGAGC	AAATCATTGA	AGCCAAAGGT	CCTGATGTTT	GTACCCCTAC	TTATTGTAAC	600
CCTAATGCCC	CTTATAGCAC	CAACACTTCA	ACCGTCGCTT	TTCAAGTGTG	GTTGAATTTT	660
GGGGTGAGAG	CCAATATCTA	CAAGCATAAT	GGCGTGGAAT	TTGGCGTGAG	AGTGCCGCTA	720
CTCATCAATA	AATTTTGTAG	CGCGGGTCCT	AACGCTACTA	ACCTTTATTA	CCATTTGAAA	780
CGGGATTATT	CGCTTTATTT	GGGGTATAAC	TACACTTTTT	AA		822

## (2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 447 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double



115

- (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...447
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84

TTGGTCCAAA	TCGTAGTCGT	GTTTTATGGT	TTGCCCCCCC	TTGGGGTCTA	TATGGATCCA	60
ATCCCGGCAG	GCATTATTGC	GTTTCTTTT	AATGTGGGGG	CATACGCTTC	AGAGACTTTG	120
AGGGCGAGCT	TTCTTTCTGT	CCCTAAAGAT	CAATGGGATT	CAAGCTTGAG	TTTGGGCTTG	180
AATTACTTGC	AAACCTTTTG	GCATGTCATC	TTTTTTC AAG	CGCTCAAAGT	CGCCACGCCA	240
AGCCTAAGTA	ACACTTTCAT	CAGCCTTTT	AAAGAACTT	CTTTAGCGTC	TGTGGTCACT	300
ATCGCAGAGG	KTTTTAG AAT	CGCACAGCAA	AAAGYGAACG	TCAGCTATGA	CTTTYGCTT	360
ATTTATTTGG	AAGKCGCTT	GATTACTGG	CTTTTTTGCT	TGGTTTTAGA	AGTGATTCAA	420
AAGCGCGTGG	AAAAAATCTT	AAATTAA				447

## (2) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 405 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...405
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85

GTGGTGGCTG	ATGAAGTTAG	GAATTTAGCT	GGGCGCACTC	AAAAGTCTTT	AGCCGAAATC	60
AATTCCTACTA	TCATGGTGAT	TGTCCAAGAA	ATCAATGATG	TGAGTTCGCA	AATGAATCTC	120
AATTTCGCAAA	AAATGGAGCG	CTTGAGCGAT	ATGAGTAAAA	GCGTGCAAGA	AACTTACGAA	180
AAAAATGAGTT	CTAATTTAAG	CTCAGTCGTT	TTAGACAGCA	ATCAAAGCAT	GGACGATTAC	240
GCTAAATCCG	GACACCAAAT	TGAAGCTATG	GTAAGCGATT	TTGCAGAAAGT	GGAAAAAGTG	300
GCTTCTAAGA	CTTTGGCTGA	TTCTTCAGAT	ATTTTAAACA	TCGCTACGCA	TGTGAGTGGA	360
ACGACCATGA	ATTTAKACAA	ACAAGTGAAT	TTGTTTAAAA	CTTAA		405

## (2) INFORMATION FOR SEQ ID NO:86:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 402 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO

116

- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...402
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86

ATGAATTACG	ATAACTATTG	GGATGAGGAC	AAACCAGAAC	TCAATATCAC	GCCTTTAGTG	60
GATGTGATGC	TTGTTTTATT	GGCTATTCTT	ATGGTAACGA	CGCCCACTCT	CACCTATAAA	120
GAAGAGATTG	CCTTGCCCTT	TGGTTCAAAA	ACTGCTAGAG	CCACTCAAGA	TAAAGTGATA	180
GAGATACGCA	TGGATAAAGA	CGCAAAAATC	TATATAGATA	GTCAAACCTA	TGAATACAMC	240
TCTTTCCCGG	ACACTTTCAA	TTTGCTTTCT	AAAAAATACG	ATAAAGATAC	TAGGGTGAGT	300
ATCCGTGCGG	ACAAGCGATT	GACCTATGAC	AAAGTGATTT	ATTTGTTAAA	AACGATTAAA	360
GAAGCGGGGT	TTTGAAAGT	TTCTTTAATC	ACAAGTCCTT	AA		402

(2) INFORMATION FOR SEQ ID NO:87:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 216 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...216
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87

ATGCCACCCA	CACSCCCCCA	AGCGAGTATT	TTAAGGCTAA	CCCTAAAAAA	CCCTTTGMGC	60
MTGCTATCTC	GTTATTCGCT	CTGTCTGTTG	AAAAAAACGC	GCTTGCAAAC	AACATCAAAC	120
AGCGCACCAA	AAGCATGCTT	GATTGCGGGC	TTATTGAAGA	AATCAAAGCC	CTTTATATTA	180
AATACCCTAA	AGATTGCGAG	CCTTTTAAAG	CCATAG			216

(2) INFORMATION FOR SEQ ID NO:88:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 654 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...654
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88

ATGCCTGTTA	TAAGAGTTTT	AGTAATGCTT	GCAACAATGA	TGATGAAATT	AGTAAAAACG	60
GCAAAAGAAA	AGAAAAGTTT	TAAGAATGTG	GGAATATCTA	TAATGGGGAT	TGCTTTTTTG	120
GAAGCGATAA	AAGACTCGAT	AAAAAAACAA	ATTAAAAAAA	GCGATTGGAT	ATGCGGGAAT	180
GTTAAGACTG	CGGATGATTA	TTTAAAAACG	CATCCTAACT	CATGGTTTAA	TTCAGCAATA	240
GGTGTAACAG	CGATAACAGC	CATGCTTATG	AATGTGTGTT	TTGCTGATGA	CCAATCCAAA	300
AAAGAAGTGG	CTCAAGCTCA	AAAGGAAGCT	GAAAACGCTA	GGGATAGAGC	GAACAAGAGT	360
GGGATAGAAC	TGGAACAAGA	AGAGCAAAAAG	ACAGAACAAG	AAAAACAAAA	GACAGAACAA	420
GAAAAACAAA	AGACAGAACA	AGAAAAACAA	AAGACAGAAC	AAGAAAAACA	AAAGACAGAA	480
CAAGAAAAAC	AAAAGACAAG	CAATATAGAG	ACTAACAATC	AAATAAAAAGT	AGAACAAGAA	540
CAACAAAAGA	CAGAACAGGA	AAAACAAAAG	ACAAACAATA	CGCAAAAAGA	TTTGGTTAAC	600
AAAGCAGAAC	AAAATTGCCA	AGAAAATCAT	AATCAATTCT	TTATTAATAA	TTAA	654

## (2) INFORMATION FOR SEQ ID NO:89:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 228 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...228
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:89

ATGGTTATTT	CTGGGCATTT	CACCACTTAT	AGCTATATTG	AGCCTTTTAT	CATTCAAATC	60
AGCCAATTTT	CTCCTGACAT	TACAACGCTA	ATGTTGTTTG	TGTTTGGGTT	AGCAGGCGTG	120
GTGGGGAGTT	TTTTGTTCGG	CCGTTTGTAT	GCGAAAAATT	CAAGAAAATT	TATCGCTTTT	180
GCAATGTTTT	TAGTCATTTG	CCCGCAACCT	CTTGCTTTTT	GTGTTTAA		228

## (2) INFORMATION FOR SEQ ID NO:90:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 576 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...576
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90

ATGAAATCTA	CAAGAATTGG	TTCTAAAATT	GTCTATGATG	TGTGTGCGGT	TGTTATTGTC	60
ATTAGCGCTG	TTATGGGCGT	TATTATCAGC	TACAAGGTTG	AAAGCGTGTT	GCAAAGCCAA	120
GCCACAGAAT	TGCTGCAAAA	AAAAGCTCAG	TTAGTCAGTT	TTAAAATTCA	AGGCATTATG	180
AAGCGCATTT	TTATGGGCGC	TAACACCCTT	GAAAGGTTTT	TAAGCGATGA	AAATGGCGCT	240
ATTAATGACA	CCCTAAAAAG	ACGCATGCTC	TCTGAGTTTT	TGTTAGCAAA	CCCTCATGTG	300

118

TTATTGGTTA	GCGCGATTTA	TACGAATAAT	AATGAACGAA	TGATCACTGC	AATGAACATG	360
GATTCAAAAA	TCGCTACCC	TAATACCGCA	CTCAATGAAA	ACATGACCMA	CCCAATCCAT	420
TCGCTCAAAA	GTATAACCCG	TTCARATCCC	TATTATAAAG	AGGTTAATGM	CRATAAAATC	480
TATRRCATRR	ACATTACCCT	CCCCCTAATR	RRCAARAATY	AAAATRTTAT	ARRRCRCWCTW	540
AATTTCKTTT	TAAACATTGA	CWGCTTTTGA	TACTGM			576

## (2) INFORMATION FOR SEQ ID NO:91:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 762 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...762
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:91

ATGGCATACA	AATATGATAG	AGACTTGGA	TTTTTAAAGC	AACTGGAATC	TAGTGATTTA	60
TTGGATTTGT	TCGAGGTGCT	TGTTTTTGGT	AAAGACGGCG	AAAAAAGACA	CAATGAAAAA	120
CTCACAAGCT	CCATAGAATA	CAAAAGGCAT	GGCGATGATT	ACGCTAAATA	CGCAGAAAGA	180
ATCGCTGAAG	AGTTGCAATA	CTATGGGAGC	AATAGTTTTG	CGAGTTTCAT	TAAAGGTGAA	240
GGAGTCTTAT	ACAAAGAGAT	TTTATGCGAT	GTGTGCGATA	AATTAAAGGT	CAATTACAAC	300
AAGAAAACCTG	AAACGACTTT	AATTGAACAA	AACATGCTTT	CTAAAATCTT	AGAAAGAAGC	360
CTAGAAGAAA	TGGATGATGA	AGAAGTGAAA	GAAATGTGCG	ATGAATTGTC	CATAAAAAAC	420
ACGGACAATT	TGAACAGACA	AGCCTTAAGC	GCGGCGACTT	TAACGCTGTT	TAAAATGGGA	480
GGCTTTAAAT	CTTATCAATT	AGCTGTCATT	GTTGCGAATG	CGGTTGCAAA	AACCATTCTA	540
GGGCGTGTTT	TATCGCTTGC	GGGCAATCAA	GTGCTTACAA	GAACTCTGAG	CTTTTTAACA	600
GGCCCTGTTG	GCTGGATCAT	TACAGGCGTA	TGGACAGCGA	TTGATATTGC	AGGGCCGGCT	660
TATAGGGTAA	CCATACCGGC	ATGCATTGTG	GTCGCCACTT	TACGCCTAAA	AACGCAACAA	720
GCCAATGAAG	ATAAGAAGTC	GTTGCAAATA	GAATCCGTTT	AA		762

## (2) INFORMATION FOR SEQ ID NO:92:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 882 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...882
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:92

TTGTTGCTTT	TTATCGTTGT	GATCACCTCT	TTGGTTAAAA	ACACGATCCC	AAATATTTGG	60
CTCACTAAAA	TCCTTTATAT	GGCTATCTTG	CTTTGCGCGA	TCGCTCATTC	TGTGGGGCYA	120
ATCTTGCGYT	GCTATGTGAG	TGGGCATTCT	CCTTGAGTA	ACGCTTATGA	GTCCATGTTT	180

119

TATATCGCAT	GGGCTTCTGT	TATCGCAGGG	TTTGTTTTAC	GAMCTAAACT	CGCGCTATCG	240
GCTTCTAGCT	TTTTGGCCGG	TATCGCGCTC	TTTGTGGCTC	ATTTAGGCTT	TATGGACCCT	300
CAAATTGGCC	CTTTAGTGCC	GGTGTAAAA	TCCTATTGGC	TCAATATCCA	TGTCTCTGTC	360
ATCACCGCTA	GTTATGGCTT	TTTGGGCTTG	TGTTTTGTGC	TAGGGATTTT	AAGTTTGGTT	420
TTGTTTATTT	TGCGCAAACA	AGGGCGTTTC	AATTTAGACA	AAACCATTCT	TTCCATTAGC	480
GCTATCAATG	AAATGAGCAT	GATTTTAGGC	CTGTTCATGC	TCACAGCCGG	GAATTTCTTA	540
GGTGGGGTGT	GGGCGAATGA	ATCTTGGGGG	CGCTATTGGG	GGTGGGACCC	TAAAGAAACT	600
TGGGCGTTGA	TTTCTATTTG	CGTCTATGCC	TTGATCTTGC	ATTTGCGTTT	TTTAGGCTCT	660
CAAAATTGGC	CCTTTATTTT	AGCGAGCAGC	AGCGTGCTAG	GGTTTTATTC	GGTTTTAATG	720
ACTTTATTTT	GGCGTGAATT	ACTACCTTTC	TGGCTTGAC	AGCTATGCCG	CAGGKSATCS	780
TTTGCCGATC	CCTACTTTTT	TATACTTTTT	GGTAGCGATA	CCTTTCGCTC	TCGTATCTTG	840
CGGTATTTCA	AACGTCATTT	GAGTTTGCCT	AAATTGGTTT	AA		882

## (2) INFORMATION FOR SEQ ID NO:93:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 429 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...429

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:93

GTGGAAATGA	TCCACACGCA	AGATTACATT	AAAATGGAAG	AAGCCGCCAC	TGAGGCGATT	60
AAGCGTAAGG	AATCTTCCAT	TTACTTGGGC	ATGGATATTT	TAAAAAATGG	GGCTGACGCT	120
TTGATTTTCAG	CGGGGCATAG	CGGGGCGACT	ATGGGTTTAG	CGACCTTGCG	TTTAGGGCGT	180
ATCAAGGGGG	TTGAAAGGCC	TGCTATTTGC	ACTTTAATGC	CTAGCGTTGG	CAAACGCCCT	240
AGCGTGCTAT	TAGACGCAGG	AGCGAACACG	GATTGCAAGC	CTGAATATTT	GATTGATTTT	300
GCTCTCATGG	GGTATGAATA	CGCTAAAAGC	GTGTTGCATT	ATGACAGCCC	TAAAGTGGGT	360
CTTTTGAGTA	ATGGTGAAGA	AGATATTAAG	GGGGGAATAC	GCTCGTTAAA	GAAACGCATA	420
AAATGTTGA						429

## (2) INFORMATION FOR SEQ ID NO:94:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 433 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...433

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94

ATGCTAGAAA	TCAAGAATTT	AAACTGCGTT	TTAAACTCTC	ATTTTTCGCT	CCAAAACATC	60
------------	------------	------------	------------	------------	------------	----

120

AATATTTCTT	TAAGTTATAG	TGAAAGGGTG	GCGATCGTGG	GCGAAAGCGG	GAGCGGGAAA	120
AGCTCTATCG	CTAATCTCGT	CATGCGATTA	AACCCTAGAT	TCAAGTCCCA	TAATGGCGAA	180
ATCCTGTTTG	AAACAACCAA	TCTTTTAAAG	GAAAGCGAAG	CGTTYMTGCA	GCATTTAAGG	240
GGGAATATTA	TCGCTTACAT	CGCCCAAGAC	CCCCTATCCA	GCCTCAACCC	CTTGCAATAA	300
ATCGGCAAGC	AAATGAGTGA	AGCCTATTTT	TTACACCATA	AAAACGCTTC	TCAAGTGTCT	360
CTTAATGAAC	AAGTTTTTAA	CGTTATGAAA	CAAGTTCAAT	TGGATGAAAA	TTTTTGGAAT	420
GTTTCTCTTA	TGT					433

## (2) INFORMATION FOR SEQ ID NO:95:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 252 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...252

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95

ATGAACTACA	AAGTTGCATC	TGCTAGAAAT	ATCGCAACGC	TTCTTTTCTT	ATTCTTTTCT	60
CAAAGTGAAG	CTTTTGATTT	GGGTAAAATC	GCTAAAATCA	AAGCGGGTGC	TGAAAGTTTC	120
TCTAAAGTCG	GTTTCAATAA	CAAACCTATC	AACAMTAATA	AAGGGATTTA	CCCTACCGAA	180
ACCTTTATGA	CGATTAATGG	CTTACATGCA	GGTGGATTTT	ACGGAGCTCT	TGCCCCAAAG	240
CGCTACGGCT	AA					252

## (2) INFORMATION FOR SEQ ID NO:96:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 393 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...393

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96

ATGGACGCTT	TAGAAATCAC	CCAAAACTC	ATCAGTTACC	CTACCATTAC	GCCCCAAGAA	60
TGCGGTATTT	TTGAATACAT	TAAATCGCTT	TTTCCTGCTT	TTAAAACCCT	AGAATGTGAA	120
AAAAATGGCG	TGAAAAACCT	TTTTTTATAC	CGCATTTTTA	ACCCCTCAA	AAAGCATGCA	180
GAAAAAGAAC	ATGCAAAAGA	AAAGCATGTR	AAAGAAAATG	TTAWGCCCTT	GCATTTTTCG	240
TTKGCAGGGC	ATATTGRTGT	CGTGCTCCT	GGGRRCAWTK	GGCRRRSKGA	TTCTTTTWWA	300
YCCATCATTA	AAGAGGGGTT	TTTATACGGT	CGTGGGGCGC	AAGACATGAA	GGGGGGCGTT	360
GGGSSGTTTT	TKAGGTGCRR	GTTWAAATTT	TAA			393

## (2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1023 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...1023
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97

ATGATTTTAA	GCATTGAAAG	TTCTTGCGAT	GACAGCTCTT	TAGCCCTTAC	AAGAATAGAG	60
GACGCCAAGC	TCATCGCTCA	TTTTAAAATC	TCTCAAGAAA	AGCACCACAG	CTCTTATGGG	120
GGCGTTGTGC	CTGAGATTGC	ATCGCGCCTG	CATGCTGAGA	ATTTGCCGCT	TTTATTAGAA	180
CGCGTTAAAA	TAAGCTTGAA	TAAGGATTTT	TCCAAAATTA	AAGCCATCGC	TATCACTAAT	240
CAGCCAGGTT	TGAGCGTTAC	TTTAATAGAG	GGTTTGATGA	TGGCAAAAGC	CTTGAGCTTG	300
TCTTTGAATT	TACCCTTGAT	TTTGGAAGAT	CATTTGAGAG	GGCATGTGTA	TTCGCTCTTT	360
ATCAATGAAA	AACAAACCCG	CATGCCTTTA	AGCGTGCTGC	TAGTCTCTGG	GGGGCATTCT	420
TTAATTTTAG	AGGCTAGAGA	TTATGAAGAC	ATTAATAATCG	TTGCCACGAG	TTTAGACGAT	480
AGCTTTGGGG	AGAGTTTGA	TAAGGTTTCA	AAAATGCTTG	ATTTAGGCTA	TCCAGGAGGC	540
CCCATAGTGG	AAAAATTAGC	CCTTGATTAT	GCACACCCAA	ACGAGCCTTT	AATGTTCCCT	600
ATCCCTTTAA	AAAACAGCCC	GAATTTGGCT	TTTAGTTTTT	CAGGTTTAAA	AAATGCGGTG	660
CGTTTGGAGG	TTGAAAAAAA	CGCCCATAA	TTGAACGATG	AGGTAAACA	AAAGATTGGC	720
TATCATTTTC	AAAGCGCGGC	TATCGAGCAT	TTAATCCAGC	AGACTAAACG	CTATTTTAAA	780
ATCAAACGCC	CTAAAATTTT	TGGCATTGTG	GGGGGAGCGA	GCCAAAATCT	AGCCTTAAGA	840
AAGGCGTTTG	AGGATTGTG	TGCTGAGTTT	GATTGCGAGC	TTGTTTTAGC	CCCTTTAGAA	900
TTTTGCAGCG	ACAATGCCGC	CATGATAGGG	CGATCAAGCC	TAGAAGCTTA	TCAAAAAAAG	960
CGCTTTATCC	CTTTAGAAAA	AGCCGATATT	TCGCCAAGAA	CGCTGTTAAA	AAATTTTGAG	1020
TGA						1023

## (2) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 507 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...507
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98

ATGTTATCTT	CTAATGATTT	GTTTATGGTC	GTTTATAGGG	CGATTTTATT	GGTGTGGGTG	60
TGCTTGGTGG	GGTATTTGTA	TCTTAAAGAA	AAAGAGTTTT	ACCATAAAAT	GAGGCGTTTA	120
GAAAAAATCT	TAGATGAATC	CTATCAAGAA	AATTATCTCT	ATTCTAAGCG	TTTGAGAGAA	180
TTAGAGGGGC	GTTTGGGAAGG	CCTTTCTTTA	GAAAAAAGCG	CTAAAGAGGA	CAGCTCATTA	240
AAAACGACTC	TTTCGCACCT	TTATAACCAG	TTGCAAGAAA	TCCAAAAATC	CATGGATAAA	300

122

GAGCGCGATT	ACTTAGAAGA	AAAAATCATT	MYTTRGAAAA	CAAWTTTWAA	GACATGGGGC	360
ATTATGCCGC	TAGCGATGAA	GTCAACGGAA	AAACAGGTTT	TGAAAATGTA	TCAAGAAGGC	420
TATAGCGTGG	ATTCTATTTT	TAAAGAATTT	AAAGTGAGTA	AGGGCGAGGT	GGAATTTATA	480
TTGAACATGG	CAGGGTTAAA	ATGGTAG				507

## (2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 366 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...366
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99

TTGGATCCCT	TTAGCCATAA	GGAGAATTTT	TTAGCCGTTG	AAACCTTTAA	AATGCTAGGC	60
AAAACAGAAA	GTAAAGACAA	TCTTAATTGG	ATGATCGCTT	TGATCATTGA	AAAAGACAAG	120
GTCTATGAGC	AAGTGGGATC	GGTGCGTTTT	GTGGTGTTG	TAGCGAGTGC	TATCATGGTG	180
TTAGCCTTAA	TCATAGCGAT	CACTCTTTTA	ATGCGAGCGA	TCGTGAGCAA	TCGTTTGGAA	240
GTCTTTTCTA	GCACCTTGTC	TCATTTCTTT	AAATTATTGA	ACAATCAAKC	CCATTCTAGC	300
RACAYTAAAT	TGGTTTRAAGC	GCGATCTAAT	GACGAATTAG	GGCGCAYGCA	AACASCTGAT	360
YAATAA						366

## (2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 450 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...450
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100

ATGGAATTTT	ATCAAGTCTA	TGACCCATTA	GGCCATATTT	GGCTGAGCGC	TTTACTCGCA	60
CTTTCGCCTA	TTGCGCTCTT	TTTTATCTCT	CTTATTGTCT	TTAAACTTAA	AGGGTATAGC	120
GCTGGGTTTT	TAAGCTTAGC	GCTTTCAATC	CTTATTGCGT	TATTTGTGTA	TAAAATGCCT	180
GTTCAAATGG	TGAGCGCGAG	TTTTTTCTAT	GGCTTTCTTT	ATGGCTTGTTG	GCCGATCGCA	240
TGGATTGTGA	TCGCTGCGAT	TTTTCTTTAC	AACCTTTTAC	TGAAGTCCGG	GTATTTTGAG	300
ATTTTAAAAG	AAAGCATTTT	AAGCTTGACG	CCGGATCATC	GCATTTTAGT	GATTTTGATC	360
GGGTTTTGTT	TTGGCTCGTT	TTTAGMMGGC	GCGRTTGGTT	TTGGAGGCCC	GCTAGCRATC	420
ACAGCGGCGA	TTTTAGTGGC	CTTGGGCTAA				450

## (2) INFORMATION FOR SEQ ID NO:101:



- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 978 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...978
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101

ATGAAAAGAA	TTTTAGTTTC	TTTGGCTGTT	TTGAGTCATA	GCGCGCATGC	TGTCAAAAC	60
CATAATTTGG	AAAGGGTGGG	AGCTTCAGGG	GTGGCTAACG	ATAAAGAAGC	GCCTTTAAGC	120
TGGAGGAGCA	AGGAAGTTAG	AAATTATATG	GGTTCTCGCA	CGGTGATTTT	TAACAAGCAA	180
CTCACTAAAA	GCGCCAATCA	AAGCATTGAA	GAAGCTTTGC	AAAATGTGCC	AGGCGTGCAT	240
ATTAGAAACT	CTACCGGTAT	TGGAGCTGTG	CCTAGCATTT	CCATTAGGGG	GTTTGGTGCT	300
GGAGGCCCCG	GGCATTCTAA	TACGGGAATG	ATTCTAGTCA	ATGGGATTCC	TATTTATGTC	360
GCGCCCTATG	TTGAAATTGG	CACGGTTATT	TTTCCTGTAA	CCTTTCAGTC	TGTGGATAGA	420
ATCAGCGTAA	CTAAGGGTGG	GGAGAGCGTG	CGTTATGGCC	CTAACGCTTT	TGGCGGTGTG	480
ATCAACATCA	TCACCAAAGG	CATTCCTACC	AATTGGGAAA	GTCAGGTGAG	CGAGAGGACC	540
ACTTTTTGGG	GCAAGTCTGA	AAACGGGGGC	TTTTTCAATC	AAAATTCTAA	AAACATTGAT	600
AAAAGCTTAG	TTAATAACAT	GCTTTTTTAA	ACCTATTTAA	GAACGGGGGG	TATGATGAAT	660
AAGCATTTTG	GAATCCAAGC	TCAAGTCAAT	TGGCTCAAAG	GGCAAGGGTT	TAGATACAAC	720
AGCCCTACGG	ATATTCAAAA	TTACATGTTA	GATTCAATTG	ATCAAATCAA	TGATAGCAAT	780
AAAATCACCG	TCTTTTTTCA	ATATTATAGT	TATTTCTTGA	CAGACCCTGG	ATCTTTAGGC	840
ATAGCCGCTT	ACAATCAAAA	TCGTTTTTCA	AACAACCGCC	CCAATAACGA	TAAAAGCGGG	900
AGAGCGAAGC	GATGGGGAGC	TGTGTATCAA	AACCTTTTTT	GGGACACGGA	TAGGGTAGGG	960
GGGGGATTTT	ACTTTTAG					978

(2) INFORMATION FOR SEQ ID NO:102:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 759 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...759
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102

TTGCGTTCAA	TTTCAAGGAT	AAAGATGCTT	TCAGTGTATG	AAAAAGGGAA	TGCCCTAGAC	60
AAAAGGGTGC	TTGAAGAATG	GCTTTTAAGC	GAAGACATTT	TAATGGAAAA	CGCCGCTATG	120
GCTTTAGAAA	GGGCGGTTTT	ACAAAACGCT	TCTTTGGGCG	CTAAGGTCAT	TATTCTTTGT	180
GGGAGTGGGG	ATAATGGAGG	TGATGGCTAT	ACTCTAGCCA	GGCGTTTAGT	GGGGCGTTTT	240
AAAACGCTGG	TCTTTGAAAT	GAAATTAGCA	AAAAGCCCCA	TGTGCCAATT	GCAAAAAGAA	300
AGGGCTAAAA	AAGTAGGGGT	AGTCATCAAA	GCATGGGAAG	AAAAGAATGA	AGATTTAGAA	360

124

TGCGATGTGT	TAGTAGATTG	CGTGGTAGGG	AGCGCTTTTA	AGGGCGGATT	AGAGCCGTTT	420
TTAGATTTTG	AAAGCCTTTC	TCAAAAAGCA	CGCTTTAAAA	TCGCTTGCGA	CATTCCTAGC	480
GGGATAGATT	CTAAAGGCAG	GGTGGATAAG	AGGGCGTTTA	ARGSCGGATA	CCGACTATCA	540
GCATGGGCGC	TATTCAAGTC	ATGCTTACTA	AGCRATAARR	CTAAARACTA	TATARRRRRAA	600
TTGAAARTRR	RRCATTTARR	RGTTTTTAAT	CAAATTTATG	AGATCCCAAC	ARACACTTTT	660
TTACTMGAAA	AAARCGATT	GAARCTGCCC	TTAAGGGATA	GAAAGAAACG	CTCACAAAGG	720
CGATTACGGG	CATGCGCATG	TGCTTTTGGG	CAAGCATAG			759

## (2) INFORMATION FOR SEQ ID NO:103:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 417 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...417
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103

ATGGCATTAG	ACAAAAGGAT	TTGGATGCAY	TTTGATCTTT	TGCCTTTTGT	GTTTATCATC	60
CCCTTGTTGG	TGGTTTCTTT	TTGTGTGATT	TTTGAGAGTA	GTGCGGTTTT	GAGCTTGAAG	120
CAAGGGGTTT	ATTATGCCAT	AGGGTTTCTT	CTCTTTTGGG	TAGTGTTTTT	TATCCCTTTC	180
AGGAAACTCG	ATCGGTGGCT	CTTTGCGCTT	TATTGGGCGT	GCGTTATTTT	ATTAGCGTTA	240
GTGGATTTTA	TGGGATCGAG	CAAGCTTGGA	GCGCAGCGAT	GGCTAGTCAT	TCCTTTCACT	300
TCTATCACCT	TACAGCCTAG	CGAGCCTGTG	AAAAATCGCY	ATTCTTTTAT	TGTTGGCGCA	360
TTTGAKYAAA	ATYAACCCGA	CCYCCTTTTA	AGGGCTATGA	TTGGGGCATG	TTTTTAA	417

## (2) INFORMATION FOR SEQ ID NO:104:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 981 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...981
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104

GTGTTAATGG	CGTTGARGCA	TAAACGCTAC	GGCTTAGAAG	CAGGGATCAA	GTATTTTACC	60
ATGGGGGCGA	TGGCGAGCGC	GTTTTTTGCT	ATGGGCGCGA	TGGCTTTTTA	CCTGCTTACA	120
GGGAGCTTGA	ATCTTGAAGT	CATTACCCTA	TACTTACACA	CTGAGGGCAT	CACAAACCCC	180
ATGCTCTTTG	CGATGGGCAC	TATTTTTTTT	ATTGGAGCGA	TTGGCTTTAA	GGTTTCTTTA	240
GTGCCTTTCC	ATACCTGGAT	GCCTGATGTG	TATGAGGGYA	ATAACCCAGT	CTTTGCGAGC	300
TATATTTCCA	TTGTGCCTAA	AATCGCTGGC	TTTGTGGTAG	CGACTCGCCT	TTTTGGGGCG	360
TTTATAGACA	CTCATACCGC	TTGGGTAGAA	GACATTTTTT	ATGTTTTGAT	CCTTATGACT	420

125

ATCACCATCC	CTAATTTTCAT	TGCTTTATGG	CAAGAAGATG	TCAAAAGGAT	GCTCGCTTAT	480
AGTTCTATTT	CGCATTCTGG	GTTCGCTTTA	GCGTGCGTGT	TTATCCACAC	TGAAGATAGC	540
CAACAAGCGA	TGTTTGTTTA	TTGGTTCATG	TTTGCCCTCA	CTTACATTGG	GGCTTTTGGC	600
CTTTTATGGC	TCTTAAAAAG	CCGGGAAAAA	ACATGGGATG	AACGCTACGA	TCACCCCTAT	660
TCTAAATTCA	ACGGCCCTTAT	CAAAACCCAC	CCCTTAGTGG	CGATCTTGGG	CGCTATTTTT	720
GTTTTTGGGC	TTGCAGGGAT	CCCGCCTTTT	AGCGTGTTTT	GGGGGAAATT	TTTAGCCGTT	780
GAAAGCGCGT	TAGAGAGCAA	TCACATTCTT	TTAGCGGTGG	TGATGTTAGT	TAATAGCGCG	840
GTGGCTGCGT	TTTATTATTT	CCGTTGGCTC	GTGGCGATGT	TTTCAATAA	GCCCTTACAA	900
ACCCAAAGCT	ACGCCAAAAC	GATATTTACA	CCCAAAACGC	CACCATGCCC	ATTTATGCGG	960
TCATTATTGC	CATGGCGTTA	G				981

## (2) INFORMATION FOR SEQ ID NO:105:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 723 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...723

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105

ATGATAAACT	CAAAGAAAAG	CTTGAAAAAG	GGCTTAAGGG	GCTTTTTTAA	AATTTTAAAG	60
GACAGAAATG	GCGCACATTT	TAGTTGCGGA	GCGACTTCAG	GGTTTGGGCT	AGAAATCGCT	120
AAGGCGTTTT	TACAAAAAAA	CCATGTGGTT	TTTGGCACAG	GGAGGCGGCA	AGAGAATTTA	180
CAAAAATTGC	AGCTCGCTTA	CCCTAAGCGT	TTCATTCCCC	TGTGTTTTGA	TCTTCAAAAC	240
AAGCCTGAAA	CTAAGCGAGC	GATAGAAACT	ATTTTTTCCA	TGACGGATCG	CATTGACGCT	300
TTAATCAATA	ACGCCGGCTT	AGCGCTAGGC	TTGAACAAGG	CTTATGAATG	CGAGTTAGAC	360
GACTGGGAAG	TCATGATAGA	CACGAATATC	AAGGGGTTGT	TGCATCTCAC	CCGCTTGATC	420
TTGCCCTCTA	TGATAGAGCA	TGACCAAGGG	ACTATCATCA	ATCTTGGTTC	TATCGCTGGC	480
ACTTACGCCT	ATCCTGGAGG	GAAGTCTAT	GGAGCGAGCA	AGGCGTTKGT	GAAACAAYYY	540
TCTYYAAATT	TGCGAGCGGA	CGTGGCTGGC	ACTAACACTA	GAGGGAGAAG	GTGGAACCCC	600
GGGTGTGTGG	CGAAACCGAA	AGTCAGCAGG	GTGCGGGGTA	AAGGCGATAA	ACCAAAGCCC	660
AAATCCGGCT	ATGAAAAACA	CCCCTTACCC	CAAACCACAA	GACAAGGGCT	AACATCGGGC	720
TAG						723

## (2) INFORMATION FOR SEQ ID NO:106:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 615 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...615

126

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:106

GTGAGCGGGG	TGGTGTTAAG	CAAATTTGAT	AGCGATTCTA	AAGGGGGTAT	CGCCTTAGGC	60
ATCACTTATC	AATTGGGCTT	ACCCTTGCGT	TTTATTGGGA	GTGGGGAAAA	AATCCCTGAT	120
TTAGACGTGT	TTATGCCTGA	AAGGATTGTG	GGGCGTTTGA	TGGGGGCTGG	AGATATTATC	180
TCGCTCGCTG	AAAAAACCGC	CAGCGTTTTA	AACCCTAATG	AAGCCAAAGA	TTTAAGCAAA	240
AAGCTCAAAA	AAGGGCAATT	CACTTTCAAC	GATTTTTTAA	ACCAAATTGA	AAAAGTGAAA	300
AAATTAGGCT	CTATGAGTTC	TCTRATCTCT	ATGATTCCAG	GTTTAGGGAA	TATGGCAAGC	360
GCGCTAAAAG	ACACGGATTT	AGAAAATTCT	TTAGAAGTGA	AAAAAATCAA	GGCCATGGTT	420
AATTCCATGA	CGAAAAAGA	GCGAGAAAAC	CCCGAGATTG	TAAACGGCAG	CCGAAGAAAA	480
AGGATCGCTT	TAGGGARCGG	CTTAGAAGYG	YCTGAAATCA	ATCGCATCAT	CAAACGCTTT	540
GATCAGGCGA	GCAAAATGGC	GAAACGACTC	ACGAATAAAA	AGGCTATTAG	CGATYTGATG	600
AATCTAWCGA	KCTAG					615

## (2) INFORMATION FOR SEQ ID NO:107:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...279

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:107

GTGGAAAAAG	CCCATCCGGA	TGTGTTTAAC	CTCTTGTTAC	AGGTTTTAGA	TGAGGGGCAT	60
TTAACCGATA	GTAAGGGCGT	GAGGGTGGAT	TTCAAAAACA	CGATTTTGAT	TTTAACCAGC	120
AATGTGGCTA	GCGGCGCGCT	TTTAGAAGAG	GATTTGAGTG	AAGCCGATAA	ACAAAAAGCG	180
ATCAAAGAGA	GCCTGAGACA	ATTCTTCAAG	CCGGAATTTT	TAAACCGCTT	AGATGAAATC	240
ATCTCCTTTA	ACGCCCTAGA	TAGTCATGCT	ATCATCTAA			279

## (2) INFORMATION FOR SEQ ID NO:108:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 246 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...246

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:108

TTGGTGT TTTT	TAGACAGGCG	TTTGATTGTG	ATGGTTACGG	ACTCTAAAGG	GAGTCGTTAT	60
ATTAATGTGC	ATATCTTATT	CCGCCAAATC	AGTTTGTATG	CGCTGTTGAG	CGTTGTGGGA	120
TCTTTATTGT	TTTAGGCGT	TTCATTACTG	GTTTTAAATA	AAGAAATTAA	AAACATTGAA	180

127

AAACAGCATG CTTTAMTCAC TAAGGAATTT GAGAAAAAAA GAGAGACGAA TGAAMAGCTT 240  
TCTTYG 246

## (2) INFORMATION FOR SEQ ID NO:109:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 702 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...702

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109

TTGAGTTTGA	TGASTGTGTT	AAATGCCAAA	GAATGCGTTY	CGCCCAT AAC	AAGAAGCGTT	60
AAGTATCATC	AGCAAAGTGC	SGAGATCAGA	GCCTTGCAAT	TACAAAGTTA	CAAAATGGCG	120
AAAATGGCGC	TAGACAATAA	CCTTAAGCTC	GTTAAAGACA	AAAAGCCAGC	CGTCATCTTG	180
GATTTAGATG	AAACCGTTTT	GAACACTTTT	GATTATGCGG	GCTATTTAGT	CAAAAAGTGC	240
ATTAAATACA	CCCCAGAAAC	TTGGGATAAA	TTTGAAAAAG	AAGGCTCTCT	TACGCTCATT	300
CCTGGAGCGC	TAGACTTTTT	AGAATACGCT	AATTCTAAGG	GCGTTAAGAT	TTTTTACATT	360
TCTAACCGCA	CCCCAAAAAA	TAAGGCATTC	ACTTTAAAAA	CGCTCAAAAG	CTTTAAGCTC	420
CCCCAAGTGA	GTGAAGAATC	CGTTTTGTTA	AAGGAAAAAG	GCAAGCCTAA	AGCCGTTAGG	480
CGGGAGTTAG	TCGCTAAGGA	TTATGCGATT	GTTTTACAAG	TGGGCGACAC	TTTGCATGAT	540
TTTGACGCCA	TTTTTGCTAA	AGACGCTAAA	AACAGCCAAG	AACAACAAGC	CAAAGTCTTG	600
CAAAACGCTC	AAAAATTTCG	CACAGAATGG	ATCATTTTAC	CCAACCTCTC	TTATGGCACA	660
TGGGAAGATG	GGCCTATAAA	AGCATGGCAA	AATAAAAAAT	AA		702

## (2) INFORMATION FOR SEQ ID NO:110:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...258

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110

ATGTTGGCGG	CTGGTTTGAC	TTTGCCTGAA	TTTGGGTGTT	ATTTAAGCCA	TTATCTTTTA	60
TGGAAAGAAT	CGTCAAAT	AGATCAACCG	GTCGTTATTT	TAGAAGATGA	TGTAACGCTA	120
GARTCTCATT	TCATGCAAGC	CTTAGAAGAT	TGTTTGAAAA	GCCCTTTTGA	TTTTGTGAGA	180
CTCTATGGGT	GTTATTGGTA	TTACCAACGA	GACAAAATTC	CATGTTTTGC	CCAAAGAATT	240
TGTATTTCTT	CCCTTTGA					258

## (2) INFORMATION FOR SEQ ID NO:111:

128

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 348 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...348
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111

TTGATCGCTT	TGAGAGTAAC	GGCTTGGAAR	GTGGYGGCCA	TGAAACGCTT	GCATTTGAGC	60
GTGAAAGACG	CTGAAACTT	TGATGCGATC	CTCAGAGAGA	GACCCTTTTT	TAAAGATTTG	120
ATAGAGTTTA	TGGTGAGTGG	TCCGGTGGTG	GTTATGGTTT	TAGAAGGCAA	AGATGCGGTG	180
GCTAAAAACA	GAGAGCTTAT	GGGAGCGACT	GATCCCAAAC	TCGCCCCAAA	AGGTACTATC	240
AGAGCGGATT	TTGCTGAGAG	CATTGACGCT	AATGCGGTGC	ATGGGAGCGA	TAGCTTGGAA	300
AACGCGCACA	ATGAAATCGC	TTTCTTTTTT	GCCGCTAGAG	AGTTTTAA		348

## (2) INFORMATION FOR SEQ ID NO:112:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1185 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...1185
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112

TTGATGTGGC	TCAAAACGCT	TACACTTCAA	ACGCTCAATA	CCGACAAAGC	CTTGCAAGAA	60
TTTTCTAAAA	CGATGGAGGC	GTTTAAAACC	AAACTCATCC	AATCCGCTAA	CGATGTGCAT	120
TCAGAGACTT	CTCGCGCCGC	TATCGCTAAC	GATTTAGAAC	GCTTAAAAGA	GCATATGATA	180
AATGTCGCTA	ATACCTCCAT	AGGGGGGGAA	TTTTTATTTG	GAGGCAGTAA	GGTGGATAGA	240
CCCCCATTTG	ATAGTAATGG	GAAATACCAT	GGCAATGGCG	AAGATTTAAA	CGCGCTTATT	300
AGCTCTGATA	ACCTTG TGCC	TTATAATATC	AGCGGGCAAG	ATTTGTTTTT	AGGCACCGAT	360
AAAGACAAAC	ACAAACTCAT	TACCACCAAC	ATTAAATTAC	TCAATCAAAA	CAAGCTCCAM	420
CCTGATGTGA	TGGACGCTTT	AGAGCATTCT	TCATTGCCTG	AAGAAGTTTT	CATTAAACCC	480
AGCGATACCT	TGCGAGAACT	CATCGGCGAT	AACGATAAAA	ACCCACCAA	TGACCCTAAA	540
GAGTTTTTTT	ATTTGCAAGG	CATTAGGCCT	GATGGCTCTA	GTTTTAAAGA	AAAATTTCGCG	600
TTGGATAAAG	CCTATCAAAA	CCAAGAAAGC	GCGACTAAAG	TGAGCGATTT	GTTGGATAAA	660
ATCGGGCATG	CTTACGGGAA	CACTTCGCAA	AATAAAGTCG	TGGATGTGAG	TTTGAACAAT	720
TGGGGGCAAA	TTGAGATTAA	AAACCTAACC	CCCGGCAGTG	AAAATTTGGA	TTTTCATTTG	780
ATTTCTAGCG	ATGGGGGATT	TGACGATTTA	GACGCCTTGC	GTTCGAGCGG	TAAAAGGGTT	840
ACTGAATATG	TCAAAAGCGC	GTTTGTAACG	GATAGGAGTT	TGAGCCAAGT	TAAAGCGGTG	900
CCTAACATGT	ATAACCCTAA	AGTGCTTGAA	ATCCCTAGCG	TGTTTGTGAC	TAAAGACAAT	960
GTTTTAGCTA	ACAAAAACAC	CAAGTTGAGC	GAGATTTTTG	GMGATAAGGT	GGAAACTTTA	1020

129

AAAATCAACG	CCAGCCGTTT	GGGCGATGAA	AGCGCTATTA	AAATCCCAAA	CCTCCCTATT	1080
AATTTGGATA	TTCCCATCT	TTTAGATGTG	AAAAACTCTA	CGATTAAAGA	TTTGAAAGAC	1140
CGATTAAAG	AACGCTTCAA	TAATGAAGGT	GGATGTGGAA	ATTGA		1185

## (2) INFORMATION FOR SEQ ID NO:113:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 309 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...309
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113

TTGAAGGCAT	TAAACGACTG	CATGGTATTT	TTTCATAAGA	AAATTATTTT	AAATTTTATC	60
TATTCTTTAA	TGGTTGCTTT	TTTATTCCAT	TTATCCTATG	GGGTTCTTTT	AAAAGCCGAT	120
GGAATGGCTA	AAAAGCAAAC	TCTTTTAGTG	GGTGAAAGGC	TTGTGTGGGA	TAAAGCTCACG	180
CTGTTAGGGT	TTTTAGAAAA	AAACCATATC	CCCCAAAAAC	TCTACTACAA	TTTGAGCTCT	240
CAAGATAAAG	AATTGAGTGC	TGAAATCCAA	AGCAATGTTA	CCTACTACMA	CTTTAAGAGA	300
TGCAAATAA						309

## (2) INFORMATION FOR SEQ ID NO:114:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1092 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...1092
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:114

ATGAAATTTT	TTCTTTTAAA	GAAATTCAGC	RAATTTTTAA	ACACTCAAAC	GCATTTTAAAC	60
CTCAAACGCT	TGAACGCGTC	TAGTTTTTTA	TTAGAGACTT	TTTCTAAAGA	AAAACACGCC	120
TTTGTTGTGG	ATTTGAGCGC	GCCTTATATT	GGTTTGTTCA	AAAAACCCCC	AGAGAGCGTT	180
TTAAAAAACA	CTTTAGCGTT	AGATTTTTGT	TTGAATAAAT	TCACCAAAAA	CGCCAAAATT	240
TTACAAGCAA	ACGTCATTGA	TAACGATCGG	ATTTTAGAAA	TCAAGGGCGC	TAAAGATTTA	300
GCTTATAAGA	GTGAACTTTT	TATTTTGCGT	TTAGAAATGA	TCCCTAAAAA	AGCCAATCTC	360
ATGATTTTAG	ATCAAGAAAA	ATGCGTGATA	GAGGCTTTTC	GTTTTAATGA	CAGGGTGGCT	420
AAAAACGATA	TTTTAGGGGC	ATTGCCTCCT	AATATTTACG	AGCATCAAGA	AGAGGATTTG	480
GATTTTAAGG	GATTGTTGGA	CATTTTAGAA	AAAGATTTTT	TATCCTATCA	GCACAAAGAA	540
TTGGAACACA	AAAAAAATCA	AATCATCAAG	CGATTAAACG	CCCCAAAAAG	ACGCTTGAAA	600
GAAAAATTAG	AAAAACTAGA	AGATCCTAAA	ACTTTACAGC	TGGAAGCGAA	AGAATTGCAA	660
ACTCAAGCCT	CATTGTTGCT	CACTTACCAG	CATTTAATCA	ACAGGCGTGA	AAATCGCGTG	720

130

```

ATTTTAAAGG ATTTTGAAGA TAAAGAATGC ATGATTGAAA TTGATAAGAG CATGCCCTTA 780
AACGCTTTTA TCAATAAAAA ATTCACTCTC AGCAAGAAAA AGAAACAAAA ATCGCAATTC 840
TTGTATTTAG AAGAAGAGAA TCTGAAAGAA AAAATCGCTT TTAAGGAAAA TCAAATCAAC 900
TATGTTAGAG ACGCTGCAGA AGAAAGCGTT TTAGAAATGT TTATGCCGGT AAAAACTCT 960
AAAATCAAAC GCCCGATGAA CGGGTATGAA GTGCTGTATT ATAAGGATTT WAAAWCGGCT 1020
TTRGAAAAAA CAAAAAGAG AATATCAAGC TTTTACAAGA CGCAARARCG AATGATTTKT 1080
GGATGCAYKT GA 1092

```

## (2) INFORMATION FOR SEQ ID NO:115:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 172 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...172

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115

```

Met Lys Gly Pro Ile Leu Trp Pro Ala Phe Ser Gln Phe Ser Asp Gln
1      5      10      15
Asp Leu Ser Asp Ile Val Ala Tyr Leu Thr Ser Ile Leu Pro Lys Asn
20      25      30
Leu Ser Asp Lys Glu Val Phe Ala Gln Ser Cys Gln Arg Cys His Ser
35      40      45
Leu Asp Tyr Ala Lys Asp Lys Ala Phe Ser Asp Pro Lys Asp Leu Ala
50      55      60
Asn Tyr Leu Gly Ser His Ala Pro Asp Leu Ser Met Met Ile Arg Ala
65      70      75      80
Lys Gly Glu His Gly Leu Asn Val Phe Ile Asn Asp Pro Gln Lys Leu
85      90      95
Leu Pro Gly Thr Ala Met Pro Arg Val Gly Leu Asn Glu Lys Ala Gln
100     105     110
Lys Gln Val Ile Ser Tyr Leu Glu Lys Ala Gly Asp Arg Lys Lys His
115     120     125
Glu Arg Asn Thr Leu Gly Ile Lys Ile Met Ile Phe Phe Ala Val Leu
130     135     140
Ser Phe Leu Ala Tyr Ala Gly Lys Glu Lys Phe Gly Ala Lys Cys Ile
145     150     155     160
Lys Phe Lys Lys Gly Gly Thr Trp Phe Tyr Asp Phe
165     170

```

## (2) INFORMATION FOR SEQ ID NO:116:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 61 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...61



131

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:116

```

Met Gln Glu Phe Ser Leu Trp Cys Asp Phe Ile Glu Arg Asp Phe Leu
1      5      10
Glu Asn Asp Phe Leu Lys Leu Ile Asn Lys Gly Ala Ile Cys Gly Xaa
      20      25      30
Thr Ser Asn Pro Ser Leu Phe Cys Glu Ala Ile Thr Lys Ser Ala Phe
      35      40      45
Tyr Gln Asp Glu Ile Ala Lys Xaa Gln Arg Gln Lys Ser
50      55      60

```

## (2) INFORMATION FOR SEQ ID NO:117:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 286 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...286

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117

```

Leu Xaa Pro Met Lys Val Ile Gln Val Phe Leu Phe Ser Asn Pro Phe
1      5      10
Cys Ala Ile Val Pro Asn Thr Glu Pro Glu His Leu Glu His Tyr Asp
      20      25      30
His Asp Leu Glu Arg Phe Phe Phe Ala Tyr Lys Tyr Phe Leu Asp His
      35      40      45
Ala Gln Lys Arg Val Ile Tyr Lys Glu Asp Pro Phe Leu Lys Asn Tyr
50      55      60
Ser Lys Asp Ala Ile Val Leu Glu Lys Lys Asp Ile Tyr Asn Ile Gln
65      70      75
Tyr Ile Leu Lys Asp Gly Glu Pro Tyr Thr Ser Phe Glu Leu Lys Asn
      85      90      95
Leu Gly Ala Phe Leu Val Trp Gly Leu Gly Glu His Asn Ala Thr Asn
      100      105      110
Ala Ser Leu Ala Ile Leu Ser Ala Leu Asp Glu Leu Asn Leu Glu Glu
      115      120      125
Ile Arg Asn Asn Xaa Leu Asn Phe Lys Gly Ile Lys Lys Arg Phe Asp
130      135      140
Ile Leu Gln Lys Asn Asn Leu Ile Leu Ile Asp Asp Tyr Ala His His
145      150      155      160
Pro Thr Glu Ile Gly Xaa Thr Leu Lys Ser Ala Arg Ile Tyr Ala Asn
      165      170      175
Leu Leu Asn Thr Gln Glu Lys Ile Ile Val Ile Trp Gln Ala His Lys
      180      185      190
Tyr Ser Arg Leu Met Asp Asn Leu Glu Glu Phe Lys Lys Cys Phe Leu
195      200      205
Glu His Cys Asp Arg Leu Ile Ile Leu Pro Val Tyr Ser Ala Ser Glu
210      215      220
Val Lys Arg Asp Ile Asp Leu Lys Ala His Phe Lys His Tyr Asn Pro
225      230      235      240
Thr Phe Ile Asp Arg Val Arg Lys Lys Gly Asp Phe Leu Glu Leu Leu
      245      250      255
Val Asn Asp Asn Val Val Glu Thr Ile Glu Lys Gly Phe Val Ile Gly
      260      265      270
Phe Gly Ala Gly Asp Ile Thr Tyr Gln Leu Arg Gly Glu Met
275      280      285

```

## (2) INFORMATION FOR SEQ ID NO:118:

132

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 61 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...61

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118

Leu	Leu	Leu	Phe	Phe	Leu	Leu	Lys	Gly	Val	Val	Phe	Ser	Leu	Gly	Phe
1				5					10					15	
Phe	Ser	Phe	Phe	Glu	Glu	Val	Ser	Gly	Ser	Phe	Xaa	Ala	Val	Ser	Leu
			20					25					30		
Xaa	Val	Leu	Ala	Leu	Val	Met	Gly	Ser	Ser	Xaa	Gly	Leu	Glu	Glu	Phe
		35				40					45				
Cys	Val	Leu	Glu	Glu	Leu	Ile	Asn	Ser	Gly	Leu	Ser	Val			
	50					55					60				

## (2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 122 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...122

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119

Met	Gly	Phe	Leu	Lys	Val	Leu	Lys	His	Asp	Ala	Leu	Gly	Gln	Val	Gly
1				5					10				15		
Asn	Ile	Val	Ile	Gly	Asn	Phe	Leu	Ile	Thr	Leu	Thr	Val	Leu	Ala	Val
			20					25					30		
Cys	Phe	Ser	Ser	Gln	Ser	Ala	Glu	Thr	Thr	Met	Leu	Thr	Leu	Ser	
		35					40				45				
Tyr	Thr	Leu	Phe	Phe	Ile	Leu	Gly	Ala	Phe	Leu	Leu	Val	Ala	Ile	Ser
	50					55				60					
Val	Gly	Ala	Ile	Lys	Asn	Leu	Asn	Ala	Leu	Phe	Ser	Lys	Arg	Gly	Val
	65				70				75					80	
Leu	Ser	Phe	Ser	Leu	Pro	Ile	Ser	Leu	Glu	Ser	Leu	Leu	Leu	Pro	Lys
				85				90					95		
Ile	Leu	Leu	Pro	Xaa	Val	Phe	Phe	Tyr	Leu	Gln	Phe	Val	Leu	Val	Cys
			100					105					110		
Gly	Glu	Arg	Ala	Phe	Gly	Leu	Leu	Pro	Phe						
		115						120							

## (2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 187 amino acids

133

(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...187

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120

```

Met Leu Lys Thr His Leu Ser Ser Ala Arg Gly Val Val Val Leu Ser
1      5      10      15
Lys Ile Leu Pro Val Asn Val Val Leu Met Val Ser Val Arg Leu Phe
20      25      30
Glu Lys Glu Leu Lys Arg Lys Pro Tyr Tyr Ile Ile Ala Ser Ala His
35      40      45
Ser Asp Glu Gly Leu Glu Lys Leu Lys Lys Xaa Gly Xaa Asp Met Val
50      55      60
Xaa Xaa Pro Thr Lys Leu Met Ala Gln Arg Val Ser Ala Asn Xaa Trp
65      70      75      80
Cys Xaa Leu Asp Met Glu Asn Ile Leu Glu Arg Phe Ile Asn Lys Lys
85      90      95
Asp Thr Leu Leu Asp Leu Glu Glu Val Ile Val Pro Lys Thr Ser Trp
100      105      110
Leu Val Leu Arg Lys Leu Lys Glu Ala His Phe Arg Glu Ile Ala Lys
115      120      125
Ala Phe Val Ile Gly Ile Thr Gln Lys Asp Gly Lys Tyr Ile Pro Met
130      135      140
Pro Asp Gly Glu Thr Ile Ile Ala Ser Glu Ser Lys Leu Leu Met Val
145      150      155      160
Gly Thr Ser Glu Gly Val Ala Thr Cys Lys Gln Leu Ile Thr Ser His
165      170      175
Gln Lys Pro Lys Glu Val Asp Tyr Ile Ser Leu
180      185

```

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 193 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...193

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121

```

Val Gly Ser Phe Leu Phe Val Gly Pro Ser Gly Val Gly Lys Thr Glu
1      5      10      15
Leu Ala Lys Glu Leu Ala Leu Asn Leu Xaa Leu His Phe Glu Arg Phe
20      25      30
Asp Met Ser Glu Tyr Lys Glu Ala His Ser Val Ala Lys Leu Ile Gly
35      40      45
Ser Pro Ser Gly Tyr Val Gly Phe Glu Gln Gly Gly Leu Leu Val Asn
50      55      60

```

134

```

Ala Ile Lys Lys His Pro His Cys Leu Leu Leu Leu Asp Glu Ile Glu
65          70          75          80
Lys Ala His Pro Asn Val Tyr Asp Leu Leu Leu Gln Val Met Xaa Asn
          85          90          95
Ala Thr Leu Ser Asp Asn Leu Gly Asn Lys Ala Ser Phe Lys His Val
          100          105          110
Ile Leu Ile Met Thr Xaa Xaa Val Gly Ser Lys Asp Lys Asp Thr Leu
          115          120          125
Gly Phe Phe Ser Thr Lys Asn Ala Lys Tyr Asp Arg Ala Val Lys Glu
          130          135          140
Leu Leu Thr Pro Glu Leu Arg Ser Arg Ile Asp Ala Ile Val Pro Phe
          145          150          155          160
Asn Ala Leu Ser Leu Glu Asp Phe Glu Thr His Cys Phe Cys Gly Ile
          165          170          175
Gly Arg Val Lys Ser Pro Ser Thr Arg Ala Arg Arg Asp Leu Lys Ile
          180          185          190
Pro

```

## (2) INFORMATION FOR SEQ ID NO:122:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 303 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...303

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:122

```

Met Ala Phe Gln Val Asn Thr Asn Ile Asn Ala Met Asn Ala His Val
1          5          10          15
Gln Ser Ala Leu Thr Gln Asn Ala Leu Lys Thr Ser Leu Glu Arg Leu
          20          25          30
Ser Ser Gly Leu Arg Ile Asn Lys Ala Ala Asp Asp Ala Ser Gly Met
          35          40          45
Thr Val Ala Asp Ser Leu Arg Ser Gln Ala Ser Ser Leu Gly Gln Ala
          50          55          60
Ile Ala Asn Thr Asn Asp Gly Met Gly Ile Ile Gln Val Ala Asp Lys
65          70          75          80
Ala Met Asp Glu Gln Leu Lys Ile Leu Asp Thr Val Lys Val Lys Ala
          85          90          95
Thr Gln Ala Ala Gln Asp Gly Gln Thr Thr Glu Ser Arg Lys Ala Ile
          100          105          110
Gln Ser Asp Ile Val Arg Leu Ile Gln Gly Leu Asp Asn Ile Gly Asn
          115          120          125
Thr Thr Thr Tyr Asn Gly Gln Ala Leu Leu Ser Gly Gln Phe Thr Asn
          130          135          140
Lys Glu Phe Gln Val Gly Ala Tyr Ser Asn Gln Ser Ile Lys Ala Ser
145          150          155          160
Ile Gly Ser Thr Thr Ser Asp Lys Ile Gly Gln Val Arg Ile Ala Thr
          165          170          175
Gly Ala Leu Ile Thr Ala Ser Gly Asp Ile Ser Leu Thr Phe Lys Gln
          180          185          190
Val Asp Gly Val Asn Asp Val Thr Leu Glu Ser Val Lys Val Ser Ser
          195          200          205
Ser Ala Gly Thr Gly Ile Gly Val Leu Ala Glu Val Ile Asn Lys Asn
          210          215          220
Ser Asn Arg Thr Gly Val Lys Ala Tyr Ala Ser Val Ile Thr Thr Ser
225          230          235          240

```

135

```

Asp Val Ala Val Gln Ser Gly Ser Leu Ser Asn Leu Thr Leu Asn Gly
                245                250                255
Ile His Leu Gly Asn Ile Ala Asp Ile Lys Xaa Asn Asp Ser Asp Gly
                260                265                270
Arg Leu Val Thr Ala Ile Asn Ala Val Thr Ser Glu Thr Gly Val Xaa
                275                280                285
Ala Tyr Thr Asp Gln Lys Gly Arg Leu Asn Leu Arg Ser Ile Gly
                290                295                300

```

## (2) INFORMATION FOR SEQ ID NO:123:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 161 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...161

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:123

```

Met Phe Phe Lys Thr Tyr Gln Lys Leu Leu Gly Ala Ser Cys Leu Ala
1      5      10      15
Leu Tyr Leu Val Gly Cys Gly Asn Gly Gly Gly Glu Ser Pro Val
      20      25      30
Glu Met Ile Xaa Asn Ser Glu Gly Thr Phe Gln Ile Asp Ser Lys Ala
      35      40      45
Asp Ser Ile Thr Ile Gln Gly Val Lys Leu Asn Arg Gly Asn Cys Ala
      50      55      60
Val Asn Phe Val Pro Val Ser Glu Thr Phe Gln Met Gly Val Leu Ser
      65      70      75      80
Gln Val Thr Pro Ile Ser Ile Gln Asp Phe Lys Asp Met Ala Ser Thr
      85      90      95
Tyr Lys Ile Phe Asp Gln Lys Lys Gly Leu Ala Asn Ile Ala Asn Lys
      100     105     110
Ile Ser Gln Leu Glu Gln Lys Gly Val Met Met Lys Pro Xaa Thr Leu
      115     120     125
Asn Phe Gly Glu Ser Leu Lys Gly Ile Ser Gln Gly Cys Asn Ile Ile
      130     135     140
Glu Ala Glu Ile Gln Thr Asp Lys Gly Ala Trp Thr Phe Asn Phe Asp
      145     150     155     160
Lys

```

## (2) INFORMATION FOR SEQ ID NO:124:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 91 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...91

SUBSTITUTE SHEET (RULE 26)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:124

Val	Cys	Leu	Gly	Leu	Ala	Asp	Val	Met	Val	Val	Leu	Ser	Leu	His	Leu
1				5				10						15	
Asn	Leu	Asn	Pro	Thr	Asn	Pro	Lys	Trp	Leu	Asn	Arg	Asp	Arg	Leu	Val
			20				25					30			
Phe	Ser	Gly	Gly	His	Ala	Ser	Ala	Leu	Val	Tyr	Ser	Leu	Leu	His	Leu
		35				40					45				
Trp	Gly	Phe	Asp	Leu	Ser	Leu	Asp	Asp	Leu	Lys	Arg	Phe	Arg	Gln	Leu
	50				55					60					
His	Ser	Lys	Thr	Pro	Gly	His	Pro	Glu	Leu	His	His	Thr	Glu	Gly	Ile
65					70					75					80
Glu	Ile	Thr	Thr	Xaa	Phe	Arg	Ala	Arg	Phe	Cys					
				85				90							

## (2) INFORMATION FOR SEQ ID NO:125:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 187 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...187

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:125

Met	Thr	Thr	Pro	Met	Ile	Ile	Ile	Ser	Leu	Glu	Met	Gly	Leu	Ser	Leu
1				5				10						15	
Val	Pro	Met	Arg	Gln	Cys	Leu	Val	Cys	Gln	Ala	Leu	Ala	Arg	Ser	Ile
			20				25					30			
Ser	Trp	Asn	Gly	Leu	Gly	Gly	Asn	Val	Arg	Asn	Thr	Lys	Val	Tyr	Gly
		35				40					45				
Lys	Phe	Ala	Ala	Tyr	His	His	Leu	Gln	Lys	Tyr	Leu	Leu	Ile	Asp	Leu
	50				55					60					
Ile	Ala	Arg	Phe	Lys	Thr	Gln	Gly	Gly	Tyr	Ile	Phe	Arg	Tyr	Asn	Thr
65				70					75						80
Asp	Asp	Tyr	Leu	Pro	Leu	Asn	Ser	Thr	Phe	Tyr	Met	Gly	Gly	Val	Thr
			85					90						95	
Thr	Val	Arg	Gly	Phe	Arg	Asn	Gly	Ser	Ile	Thr	Pro	Lys	Asp	Glu	Phe
			100				105						110		
Gly	Leu	Trp	Leu	Gly	Gly	Asp	Gly	Ile	Phe	Thr	Xaa	Ser	Thr	Glu	Leu
		115				120						125			
Ser	Tyr	Gly	Val	Leu	Lys	Ala	Ala	Lys	Met	Arg	Leu	Ala	Trp	Phe	Phe
	130				135					140					
Asp	Phe	Gly	Phe	Leu	Thr	Phe	Xaa	Thr	Pro	Thr	Arg	Gly	Ser	Phe	Phe
145				150					155						160
Tyr	Asn	Ala	Xaa	Thr	Thr	Thr	Ala	Asn	Phe	Lys	Asp	Tyr	Xaa	Val	Val
			165					170						175	
Gly	Xaa	Xaa	Phe	Glu	Xaa	Ala	Thr	Trp	Arg	Ala					
			180				185								

## (2) INFORMATION FOR SEQ ID NO:126:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

137

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...104

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126

```

Met Gln Ala Leu Lys Ser Leu Leu Glu Val Ile Thr Lys Leu Gln Asn
1      5      10      15
Leu Gly Gly Tyr Leu Met His Ile Ala Ile Phe Ile Ile Phe Ile Trp
      20      25      30
Ile Gly Gly Leu Lys Phe Val Pro Tyr Glu Ala Glu Gly Ile Ala Pro
      35      40      45
Phe Val Xaa Asn Ser Pro Phe Phe Ser Phe Met Tyr Lys Phe Glu Lys
      50      55      60
Pro Ala Tyr Lys Gln His Lys Met Ser Glu Ser Gln Ser Met Gln Glu
      65      70      75      80
Glu Met Gln Asp Asn Pro Lys Ile Val Glu Asn Lys Xaa Trp His Lys
      85      90      95
Glu Asn Arg Thr Ser Phe Ser Gly
      100

```

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 182 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...182

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127

```

Met Gln Phe Glu Glu Met Lys Glu Leu Ala His Gln Ile Gly Val Phe
1      5      10      15
Tyr His Val Gly Val Asp Gly Ile Ala Leu Phe Leu Leu Leu Asn
      20      25      30
Ala Ile Val Val Leu Leu Ser Val Val Tyr Val Lys Glu Arg Arg Lys
      35      40      45
Asp Phe Val Ile Cys Leu Leu Leu Xaa Gly Ile Leu Met Gly Val
      50      55      60
Phe Ser Ser Leu Asn Val Ile Phe Phe Tyr Ala Phe Trp Glu Ile Ser
      65      70      75      80
Leu Leu Pro Val Leu Tyr Leu Ile Gly Arg Phe Gly Arg Asn Asn Lys
      85      90      95
Ile Tyr Ser Gly Met Lys Phe Phe Leu Tyr Thr Phe Leu Ala Ser Leu
      100      105      110
Cys Met Leu Leu Gly Ile Leu Tyr Ile Gly Tyr Asp Tyr Ala Asn Asn
      115      120      125
Tyr Gly Met Met Ser Phe Asp Ile Leu Asp Trp Tyr Gln Leu Asn Phe
      130      135      140
Ser Ser Gly Ile Lys Thr Trp Leu Phe Val Ala Phe Leu Ile Gly Ile
      145      150      155      160
Ala Val Lys Ile Pro Leu Phe Pro Phe Thr His Gly Cys Leu Met Arg
      165      170      175

```

SUBSTITUTE SHEET (RULE 26)

138

Ile Leu Thr Pro Pro Leu  
180

## (2) INFORMATION FOR SEQ ID NO:128:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 116 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...116
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:128

Val	Lys	Lys	Tyr	Ala	Glu	Asp	Phe	Ile	Thr	Lys	Asp	Glu	Val	Lys	Ser
1				5					10					15	
Leu	Leu	Glu	Arg	Leu	Ala	Lys	Asp	Tyr	Pro	Thr	Ile	Val	Glu	Glu	Ser
			20					25					30		
Lys	Lys	Ile	Pro	Thr	Gly	Ala	Ile	Arg	Ser	Val	Leu	Gln	Ala	Leu	Leu
		35					40					45			
His	Glu	Lys	Ile	Pro	Ile	Lys	Asp	Met	Leu	Thr	Ile	Leu	Glu	Thr	Ile
	50					55				60					
Thr	Asp	Ile	Ala	Pro	Leu	Val	Gln	Asn	Asp	Val	Asn	Ile	Leu	Thr	Glu
65					70				75					80	
Gln	Val	Arg	Ala	Arg	Leu	Ser	Arg	Val	Ile	Thr	Asn	Ala	Phe	Lys	Ser
			85					90					95		
Glu	Asp	Gly	Arg	Leu	Lys	Phe	Leu	Thr	Phe	Ser	Thr	Asp	Xaa	Glu	Gln
			100				105						110		
Phe	Xaa	Ala	Gln												
			115												

## (2) INFORMATION FOR SEQ ID NO:129:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 240 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...240
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:129

Met	Met	Lys	Asn	Lys	Arg	Ser	Gln	Asn	Ser	Pro	Tyr	Val	Thr	Pro	Asp
1				5					10					15	
Asn	Pro	Tyr	Leu	Thr	Leu	Glu	Lys	Ala	Leu	Gly	Tyr	Ser	Phe	Lys	Asp
			20					25					30		
Lys	Arg	Leu	Glu	Gln	Ala	Leu	Thr	His	Lys	Ser	Cys	Lys	Leu	Ala	
		35				40					45				
Leu	Asn	Asn	Glu	Arg	Leu	Glu	Phe	Leu	Gly	Asp	Ala	Val	Leu	Gly	Leu
	50					55				60					
Val	Ile	Gly	Glu	Leu	Leu	Tyr	His	Lys	Phe	Xaa	Xaa	Xaa	Asp	Gly	Gly



139

65				70					75				80		
Lys	Leu	Ser	Lys	Leu	Arg	Ala	Ser	Ile	Val	Ser	Ala	His	Gly	Phe	Thr
				85					90					95	
Lys	Leu	Ala	Lys	Ala	Ile	Ala	Leu	Gln	Asp	Tyr	Leu	Arg	Val	Ser	Ser
			100					105					110		
Ser	Glu	Glu	Ile	Ser	Lys	Gly	Arg	Glu	Lys	Pro	Ser	Ile	Leu	Ser	Ser
		115					120					125			
Ala	Phe	Glu	Ala	Leu	Met	Ala	Gly	Val	Tyr	Leu	Glu	Ala	Gly	Leu	Ala
	130					135					140				
Lys	Val	Arg	Lys	Ile	Ile	Gln	Asn	Leu	Leu	Asn	Arg	Ala	Tyr	Lys	Arg
145				150						155					160
Leu	Asp	Leu	Glu	His	Leu	Phe	Met	Asp	Tyr	Lys	Thr	Ala	Leu	Gln	Glu
			165					170						175	
Leu	Thr	Gln	Xaa	Gln	Phe	Cys	Val	Ile	Pro	Thr	Tyr	Gln	Leu	Leu	Gln
		180						185					190		
Glu	Lys	Gly	Pro	Asp	His	His	Lys	Glu	Phe	Glu	Met	Ala	Leu	Tyr	Ile
	195						200					205			
Gln	Asp	Lys	Met	Tyr	Ala	Thr	Ala	Lys	Gly	Lys	Ser	Lys	Lys	Glu	Ala
	210					215					220				
Glu	Gln	Gln	Cys	Ala	Tyr	Gln	Ala	Leu	Gln	Asn	Leu	Arg	Lys	Pro	Asn
225					230					235					240

## (2) INFORMATION FOR SEQ ID NO:130:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 228 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...228

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:130

Leu	Leu	Val	Leu	Leu	Asn	Leu	Lys	Xaa	Thr	Pro	Asn	Leu	Met	Trp	Pro
1				5					10					15	
Leu	Asp	Ile	Ile	Val	Val	Val	Ala	Trp	Val	Leu	Trp	Gly	Val	Asn	Met
			20					25					30		
Phe	Gly	Ser	Met	Ser	Val	Arg	Arg	Glu	Asn	Thr	Ile	Tyr	Val	Ser	Leu
	35						40					45			
Trp	Tyr	Tyr	Ile	Ala	Thr	Tyr	Val	Gly	Ile	Ala	Val	Met	Tyr	Ile	Phe
	50					55					60				
Asn	Asn	Leu	Ser	Ile	Pro	Thr	Tyr	Phe	Val	Ala	Asp	Met	Gly	Ser	Val
65				70					75					80	
Trp	His	Xaa	Ile	Ser	Met	Tyr	Ser	Gly	Ser	Asn	Asp	Ala	Leu	Ile	Gln
			85					90						95	
Trp	Trp	Trp	Gly	His	Asn	Ala	Val	Ala	Phe	Val	Phe	Thr	Ser	Gly	Val
			100					105					110		
Ile	Gly	Thr	Ile	Tyr	Tyr	Phe	Leu	Pro	Lys	Glu	Ser	Gly	Gln	Pro	Ile
	115						120					125			
Phe	Ser	Tyr	Lys	Leu	Thr	Leu	Phe	Ser	Phe	Trp	Ser	Leu	Met	Phe	Val
	130					135					140				
Tyr	Ile	Trp	Ala	Gly	Gly	His	His	Leu	Ile	Tyr	Ser	Thr	Val	Xaa	Asp
145				150						155					160
Xaa	Val	Gln	Thr	Leu	Ser	Ser	Xaa	Phe	Ser	Val	Val	Leu	Ile	Leu	Pro
			165					170						175	
Ser	Xaa	Gly	Thr	Ala	Ile	Asn	Met	Leu	Leu	Xaa	Met	Arg	Gly	Gln	Trp
		180					185						190		
His	Gln	Xaa	Lys	Glu	Ser	Pro	Leu	Ile	Lys	Phe	Leu	Val	Leu	Ala	Ser
	195					200						205			
Thr	Phe	Tyr	Met	Leu	Ser	Thr	Leu	Glu	Gly	Ser	Ile	Gln	Ala	Ile	Lys

SUBSTITUTE SHEET (RULE 26)

140

```

      210                      215                      220
Ser Val Asn Ala
225

```

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 162 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...162

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131

[illegible]

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...59

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132

Met Ala Phe Ile Leu Thr Thr Asn Leu Phe Ile Lys Ser Phe Thr Asn  
1 5 10 15

141

```

Ser Ile Arg Ile Thr Gly Cys Ile Ile Ser Pro Asn Val Phe Phe Ala
                20                25                30
Tyr Glu Phe Cys Ala Leu Gly Phe Arg Lys Gly Gly Leu Ile Leu Asp
                35                40                45
Asn Phe Ser Lys Phe Val Ser His Arg Leu Gln
                50                55

```

## (2) INFORMATION FOR SEQ ID NO:133:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 248 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...248

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:133

```

Val Arg Phe Phe Ile Phe Leu Ile Leu Ile Cys Pro Leu Ile Cys Pro
1          5          10          15
Leu Met Ser Ala Asp Ser Ala Leu Pro Ser Val Asn Leu Ser Leu Asn
                20                25                30
Ala Pro Ser Asp Pro Lys Gln Leu Val Thr Thr Leu Asn Val Ile Ala
                35                40                45
Leu Leu Thr Leu Leu Val Leu Ala Pro Ser Leu Ile Leu Val Met Thr
                50                55                60
Ser Phe Thr Arg Leu Ile Val Val Phe Ser Phe Leu Arg Thr Ala Leu
                65                70                75                80
Gly Thr Gln Gln Thr Pro Pro Thr Gln Ile Leu Val Ser Leu Ser Leu
                85                90                95
Ile Leu Thr Phe Phe Ile Met Glu Pro Ser Leu Lys Lys Ala Tyr Asp
                100                105                110
Thr Gly Ile Lys Pro Tyr Met Asp Lys Lys Ile Ser Tyr Thr Glu Ala
                115                120                125
Phe Glu Lys Ser Thr Leu Pro Phe Lys Glu Phe Met Leu Lys Asn Thr
                130                135                140
Arg Glu Lys Asp Leu Ala Leu Phe Phe Arg Ile Arg Asn Leu Pro Asn
                145                150                155                160
Pro Lys Thr Pro Asp Asp Val Ser Leu Ser Val Leu Ile Pro Ala Phe
                165                170                175
Met Ile Ser Glu Leu Lys Thr Ala Phe Gln Ile Gly Phe Leu Leu Tyr
                180                185                190
Leu Pro Phe Leu Val Ile Asp Met Val Ile Ser Ser Ile Leu Met Ala
                195                200                205
Met Gly Met Met Met Leu Pro Pro Val Met Ile Ser Leu Pro Phe Lys
                210                215                220
Ile Leu Val Phe Ile Leu Val Asp Gly Phe Asn Leu Leu Thr Glu Asn
                225                230                235                240
Leu Val Ala Ser Phe Lys Met Val
                245

```

## (2) INFORMATION FOR SEQ ID NO:134:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

SUBSTITUTE SHEET (RULE 26)

142

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...166

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134

```

Leu Leu Val Thr Phe Leu Asn Gly Phe Asp Pro Lys Ile Ala Asn Leu
1      5      10
Arg Lys Ala Cys Asn Val Tyr Ser Val Gly Val Ile Tyr Ile Val Thr
      20      25      30
Thr Asn Thr Leu Asn Ile Leu Ser Cys Glu Ser Phe Glu Ile Leu Glu
      35      40      45
Lys Arg Glu Leu Asp Thr Ser Gly Val Thr Lys Thr Ser Thr Pro Phe
      50      55      60
Phe Ser Arg Val Glu Gly Ile Asp Ala Gly Thr Leu Gly Lys Leu Phe
65      70      75      80
Ser Gly Ser Gln Ser Lys Asn Tyr Phe Ala Tyr Tyr Asp Ala Leu Val
      85      90      95
Lys Lys Glu Lys Arg Lys Glu Val Arg Ile Glu Lys Lys Glu Glu Arg
      100      105      110
Ile Asp Ala Arg Glu Asn Lys Arg Glu Ile Lys Gln Glu Ala Ile Lys
      115      120      125
Glu Pro Lys Lys Ala Asn Gln Gly Thr Glu Asn Ala Pro Thr Leu Glu
      130      135      140
Glu Lys Xaa Tyr Gln Lys Ala Glu Arg Lys Phe Asp Ala Lys Xaa Xaa
145      150      155      160
Arg Arg Ser Phe Lys Xaa
      165

```

(2) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 127 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...127

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135

```

Met Glu Asn Ser Thr Leu Tyr Ile Val Ile Ala Gly Leu Trp Leu Ala
1      5      10      15
Val Gly Phe Gly Ile Phe Leu Lys Lys Leu Asp Met Pro Val Ile Ile
      20      25      30
Gly Tyr Ile Cys Thr Gly Thr Val Leu Ala Ala Phe Phe Lys Ile Asn
      35      40      45
Asp Phe Asn Leu Leu Ser Asp Ile Gly Glu Phe Gly Ile Val Phe Leu
      50      55      60
Met Phe Met Ile Gly Ile Glu Phe Asn Phe Asp Lys Leu Lys Ser Ile
      65      70      75      80
Lys Gln Glu Val Leu Val Phe Gly Leu Leu Gln Val Val Leu Cys Ala
      85      90      95
Leu Ile Ala Phe Leu Leu Gly Tyr Phe Val Leu Gly Leu Ser Pro Ile
      100      105      110
Phe Ser Leu Val Leu Gly Met Gly Leu Ser Leu Ser Ser Thr Ala

```

SUBSTITUTE SHEET (RULE 26)

143

115

120

125

## (2) INFORMATION FOR SEQ ID NO:136:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...16

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:136

Leu	Leu	Leu	Met	Leu	Asn	Lys	Pro	Lys	Pro	Leu	Phe	Leu	Xaa	Leu	Gly
1				5					10					15	

## (2) INFORMATION FOR SEQ ID NO:137:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...350

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:137

Met	Ala	Leu	Arg	Val	Leu	Leu	Phe	Phe	Cys	Phe	Leu	Phe	Leu	Gln	Ala
1				5					10					15	
Glu	Asp	Lys	Ser	Gln	Glu	Leu	Ser	Ser	Ile	Gln	Lys	Gln	Met	Ala	Leu
		20						25					30		
Val	Asp	Lys	Lys	Leu	Ala	Lys	Asp	Asp	Asn	Val	Trp	Leu	Lys	Lys	Phe
		35					40					45			
Glu	Asn	Tyr	Lys	Ile	Tyr	Asn	Gln	Ile	Tyr	Thr	Glu	Lys	Glu	Ser	Val
	50				55						60				
Arg	Gln	Glu	Leu	Arg	Arg	Leu	Lys	Asn	Lys	Lys	Ser	Lys	Asp	Leu	Leu
65				70					75					80	
Lys	Ile	Ser	Thr	Leu	Glu	His	Thr	Leu	Lys	Ala	Leu	Glu	Ser	Gln	Gln
			85					90					95		
Lys	Met	Phe	Glu	Ser	Tyr	Gly	Val	Asn	Pro	Phe	Lys	Asp	Leu	Ile	Glu
		100					105						110		
Arg	Pro	Asn	Ile	Pro	Asn	Ile	Pro	Asn	Ile	Ala	Asn	Pro	Ile	Ala	Ile
	115				120						125				
Ile	Asp	Gly	Ile	Ser	Phe	Ile	Lys	Ser	Met	Arg	Leu	Lys	His	Glu	Asn
	130				135					140					
Leu	Lys	Asn	Asn	Gln	Thr	Ser	Leu	Gly	Glu	Val	Leu	Lys	Leu	Leu	Asp
145				150					155					160	
Gln	Lys	His	Gln	Leu	Leu	Asn	Gln	Trp	His	Ala	Leu	Asp	Lys	Ser	Ala
			165					170					175		
Lys	Leu	Ser	Asp	Glu	Ile	Tyr	Gln	Thr	Gln	Ala	Lys	Arg	Leu	Glu	Leu
			180					185					190		

144

Gln Gly Ala Gln Asn Ile Leu Lys Thr Thr Ile Gly Ile Phe Gln Lys  
 195 200 205  
 Asp Ser Asp Glu Ala Ile Ser Ile Val Lys Ser Gln Val Lys Asn Gln  
 210 215 220  
 Leu Phe Lys Leu Val Tyr Val Phe Leu Ala Ala Leu Leu Ser Val Val  
 225 230 235 240  
 Phe Ala Trp Ile Leu Lys Ile Ile Ser Ser Lys Tyr Ile Glu Asn Asn  
 245 250 255  
 Glu Arg Val Tyr Thr Val Asn Lys Ala Ile Asn Phe Val Asn Val Ser  
 260 265 270  
 Val Ile Xaa Xaa Ile Xaa Leu Phe Ser Tyr Leu Glu Asn Val Thr Tyr  
 275 280 285  
 Leu Val Thr Val Leu Gly Phe Ala Ser Ala Gly Leu Ala Ile Xaa Met  
 290 295 300  
 Lys Asp Leu Phe Met Ser Leu Leu Gly Trp Phe Ile Ile Leu Ile Gly  
 305 310 315 320  
 Gly Ser Val His Val Gly Asp Arg Val Arg Ile Ala Lys Gly Thr Asp  
 325 330 335  
 Ile Phe Ile Gly Asp Val Leu Asp Thr Ser Asn Val Val His  
 340 345 350

## (2) INFORMATION FOR SEQ ID NO:138:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 99 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...99

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138

Met Lys Glu Gln Glu Trp Asp Leu Ser Ala Leu Phe Glu Asn Lys Glu  
 1 5 10 15  
 Ser Ala Glu Glu Phe Leu Lys Thr Leu Gln Thr Glu Val Gln Glu Phe  
 20 25 30  
 Glu Asn Ala Tyr Gln Asn Asn Leu Lys Asn Leu Asp Ala Ala Lys Phe  
 35 40 45  
 Ala Asn Thr Leu Lys His Tyr Glu Asn Leu Ser Glu Lys Ile Ser Arg  
 50 55 60  
 Ala Met Ala Tyr Ala Asn Tyr Phe Leu Pro Arg Thr Leu Lys Lys Arg  
 65 70 75 80  
 Ser Phe Ile Arg Asn Ala Asn Gly Leu Cys Lys Tyr Pro Thr Thr Pro  
 85 90 95  
 Phe Ile Leu

## (2) INFORMATION FOR SEQ ID NO:139:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

145

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...78

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:139

```

Leu Arg Val Gly Met Tyr Glu Val Cys Asn His Lys Asp Gly Thr Ala
1          5          10          15
Tyr His Ser Thr Arg Gly Ser Lys Val Thr Leu Ala Cys Lys Thr Gly
          20          25          30
Thr Ala Gln Val Val Glu Ile Ala Gln Asn Ile Val Asn Arg Met Lys
          35          40          45
Glu Lys Asp Met Glu Tyr Phe His Xaa Ser His Xaa Trp Ile Thr Xaa
          50          55          60
Tyr Leu Xaa Pro Met Lys Asn Pro Asn Thr Leu Ser Leu Phe
65          70          75

```

## (2) INFORMATION FOR SEQ ID NO:140:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 52 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...52

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:140

```

Leu Gly Leu Val Xaa Gly Ile Ser Leu Leu His Leu Ser Leu Glu Gln
1          5          10          15
Lys Ile Ser Val Phe Leu Gly Xaa Asn Leu Met Leu Tyr Pro Val Xaa
          20          25          30
Glu Val Leu Phe Ser Ile Leu Arg Arg Lys Ile Lys Arg Gln Lys Ala
          35          40          45
Thr His Ala Gly
50

```

## (2) INFORMATION FOR SEQ ID NO:141:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 377 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM:
- Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...377

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:141

```

Leu Ala Gln Pro Val Gln Val Arg Thr Val Phe Met Ser Met Thr Leu
1          5          10          15

```

146

```

Asn Ala Met Gly Gln Phe Ala Tyr Asn Phe Pro Ala Asn Ile Ser Lys
      20      25      30
Asp Lys Gln Lys Leu Thr Met Val Tyr Met Asp Lys Asp Tyr Gly Ala
      35      40      45
Tyr Gly Asn Ile Val Ala Met Gly Gly Glu Tyr Val Lys Ile Glu Leu
      50      55      60
Gly Thr Asp Thr Gly Leu Asn Pro Phe Ala Trp Ala Ala Cys Val Gln
      65      70      75      80
Lys Thr Asn Ala Thr Met Glu Gln Lys Gln Thr Ala Ile Ser Val Val
      85      90      95
Lys Glu Leu Val Lys Asn Leu Ala Thr Lys Ser Asp Glu Lys Asp Glu
      100      105      110
Asn Gly Asn Ser Ile Ser Phe Ser Leu Ala Asp Ser Asn Thr Leu Ala
      115      120      125
Ala Ala Val Thr Asn Leu Ile Thr Gly Asp Met Asn Leu Asp Tyr Pro
      130      135      140
Ile Thr Gln Leu Ile Asn Ala Phe Gly Lys Asp His Asn Asp Pro Asn
      145      150      155      160
Gly Leu Val Ala Arg Leu Ala Pro Phe Cys Lys Ser Thr Asn Gly Glu
      165      170      175
Phe Gln Trp Leu Phe Asp Asn Lys Ala Thr Asp Arg Leu Asp Phe Ser
      180      185      190
Lys Thr Ile Ile Gly Val Asp Gly Ser Ser Phe Leu Asp Asn Asn Asp
      195      200      205
Val Ser Pro Phe Ile Cys Phe Tyr Leu Phe Ala Arg Ile Gln Glu Ala
      210      215      220
Met Asp Gly Arg Arg Phe Val Leu Asp Ile Asp Glu Ala Trp Lys Tyr
      225      230      235      240
Leu Gly Asp Pro Lys Val Ala Tyr Phe Val Arg Asp Met Leu Lys Thr
      245      250      255
Ala Arg Lys Arg Asn Ala Ile Val Arg Leu Ala Thr Gln Ser Ile Thr
      260      265      270
Asp Leu Leu Ala Cys Pro Ile Ala Asp Thr Ile Arg Glu Gln Cys Pro
      275      280      285
Thr Lys Ile Phe Leu Arg Asn Asp Gly Gly Asn Leu Ser Asp Tyr Gln
      290      295      300
Arg Leu Ala Asn Val Thr Glu Lys Glu Phe Glu Ile Ile Thr Lys Gly
      305      310      315      320
Leu Asp Arg Lys Ile Leu Tyr Lys Gln Asp Gly Ser Pro Ser Val Ile
      325      330      335
Ala Ser Phe Asn Leu Arg Gly Ile Pro Lys Glu Tyr Leu Lys Ile Leu
      340      345      350
Ser Thr Asp Thr Val Phe Val Lys Glu Ile Asp Lys Ile Ile Gln Asn
      355      360      365
His Ser Ile Ile Asp Lys Tyr Gln Pro
      370      375

```

## (2) INFORMATION FOR SEQ ID NO:142:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...154

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:142

```

Met Lys Leu Val Ser Leu Ile Val Ala Leu Val Phe Cys Cys Phe Leu
 1              5              10              15

```



147

```

Gly Ala Val Glu Leu Pro Gly Val Tyr Gln Thr Gln Glu Phe Leu Tyr
      20      25      30
Met Lys Ser Ser Phe Val Glu Phe Glu His Asn Gly Lys Phe Tyr
      35      40      45
Ala Tyr Gly Ile Ser Asp Val Xaa Xaa Ser Lys Ala Lys Lys Asp Lys
      50      55      60
Leu Asn Pro Asn Pro Lys Leu Arg Asn Arg Ser Asp Lys Gly Val Val
      65      70      75      80
Phe Leu Ser Asp Leu Ile Lys Val Gly Glu Gln Ser Tyr Lys Gly Gly
      85      90      95
Lys Ala Xaa Asn Phe Xaa Asp Gly Lys Thr Xaa His Val Arg Val Thr
      100      105      110
Gln Xaa Ser Asn Gly Asp Leu Xaa Phe Thr Ser Ser Tyr Xaa Lys Trp
      115      120      125
Gly Tyr Val Gly Lys Thr Phe Thr Trp Lys Arg Leu Ser Asp Glu Glu
      130      135      140
Ile Lys Asn Leu Lys Leu Lys Arg Phe Asn
      145      150

```

## (2) INFORMATION FOR SEQ ID NO:143:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143

```

Leu Glu Thr Leu Phe Leu Val
1          5

```

## (2) INFORMATION FOR SEQ ID NO:144:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 114 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...114

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144

```

Met Xaa Thr His Asp Arg Arg Lys Leu Arg Ile Xaa Leu Thr Gln Thr
1          5          10          15
Thr Thr Leu Val Ala Thr Ile Gly Ser Asn Ala Pro Tyr Ile Gly Leu
      20      25      30
Leu Gly Thr Val Met Gly Ile Met Leu Thr Phe Met Asp Leu Gly Ser
      35      40      45
Ala Ser Gly Ile Asp Thr Lys Ala Ile Met Thr Asn Leu Ala Leu Ala

```

148

```

      50              55              60
Leu Lys Ala Thr Gly Met Gly Leu Leu Val Ala Ile Pro Ala Ile Val
65              70              75              80
Ile Tyr Asn Leu Leu Val Arg Lys Ser Glu Ile Leu Val Thr Lys Trp
      85              90
Asp Ile Phe His His Pro Val Asp Thr Gln Ser His Glu Val Tyr Ser
      100              105              110
Lys Ala

```

## (2) INFORMATION FOR SEQ ID NO:145:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 67 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...67

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145

```

Met Gln Asp Leu Asp Asn Asn Met Ser Leu Asp Thr Ala His Asn Thr
1              5              10              15
Leu Ser Ser Asn Gly Lys Asn Ile Thr Ile Ala Gly Val Val Lys Ala
      20              25              30
Leu Gln Lys Ile Gly Val Ser Ala Lys Gly Met Val Ser Ile Leu Gln
      35              40              45
Ala Leu Lys Lys Ser Gly Ala Ile Ser Ala Lys Trp Arg Tyr Tyr Asp
      50              55              60
Lys Gln Gln
65

```

## (2) INFORMATION FOR SEQ ID NO:146:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 88 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...88

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146

```

Leu His Pro Leu Ala Asp Val Phe Val Val Asn Asp Lys Arg Xaa Val
1              5              10              15
Leu Ala Met Val Xaa Met Leu Ile Xaa Ser Leu Ala Asn Ile Phe Phe
      20              25              30
Asn Tyr Leu Phe Ile Phe Xaa Leu Glu Val Gly Val Gln Gly Xaa Ala
      35              40              45
Ile Val Thr Val Ile Gly His Ala Ile Gly Gly Leu Val Leu Met Gln
      50              55              60

```

149

His Phe Trp Arg Lys Lys Gly Glu Leu Tyr Phe Ile Lys Leu Ile Phe  
 65 70 75 80  
 Phe Ile Phe Ser His Phe Phe Ser  
 85

## (2) INFORMATION FOR SEQ ID NO:147:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 276 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...276

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:147

Met Leu Arg Lys Asn Ile Leu Ala Tyr Tyr Gly Ala Asn Phe Leu Leu  
 1 5 10 15  
 Ile Ile Ala Gln Ser Leu Pro His Ala Ile Leu Thr Pro Leu Leu Leu  
 20 25 30  
 Ser Lys Gly Leu Ser Leu Ser Glu Ile Leu Leu Val Gln Thr Phe Phe  
 35 40 45  
 Ser Phe Cys Val Leu Val Ala Glu Tyr Pro Ser Gly Val Leu Ala Asp  
 50 55 60  
 Leu Met Ser Arg Lys Asn Leu Phe Leu Val Ser Asn Val Phe Leu Ile  
 65 70 75 80  
 Ala Ser Phe Ser Leu Val Leu Phe Phe Asp Ser Phe Ile Leu Met Leu  
 85 90 95  
 Leu Ala Trp Gly Leu Tyr Gly Leu Tyr Ser Ala Cys Ser Ser Gly Thr  
 100 105 110  
 Ile Glu Ala Ser Leu Ile Thr Asp Ile Lys Glu Asn Lys Lys Asp Leu  
 115 120 125  
 Ser Lys Phe Leu Ala Lys Asn Asn Gln Ile Thr Tyr Leu Gly Met Ile  
 130 135 140  
 Ile Gly Ser Ser Leu Gly Ser Phe Leu Tyr Leu Lys Val His Ala Met  
 145 150 155 160  
 Leu Tyr Val Val Gly Ile Phe Leu Ile Met Leu Cys Ala Leu Thr Ile  
 165 170 175  
 Ile Ile Tyr Phe Lys Glu Lys Glu Gly Asp Phe Lys Ser Gln Lys Asn  
 180 185 190  
 Leu Lys Leu Leu Lys Glu Gln Val Lys Gly Ser Leu Lys Glu Leu Lys  
 195 200 205  
 Asp Asn Pro Lys Leu Lys Ile Leu Leu Val Gly His Leu Ile Thr Pro  
 210 215 220  
 Val Phe Phe Met Ser His Phe Gln Met Trp Gln Ala Tyr Phe Leu Lys  
 225 230 235 240  
 Gln Gly Val Lys Glu Gln Tyr Leu Phe Val Phe Tyr Ile Ala Phe Gln  
 245 250 255  
 Val Ile Ser Ile Pro His Ser Phe Phe Lys Ser Gln Lys Leu Xaa Ala  
 260 265 270  
 Lys Lys Ser Pro  
 275

## (2) INFORMATION FOR SEQ ID NO:148:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 93 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

150

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...93

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148

```

Leu Tyr Pro Pro Gly Ser Val Val Lys Met Gly Val Gly Leu Ser Phe
1      5      10      15
Leu Glu Asn Leu His Ile Thr Glu Asn Thr Thr Ile Pro Thr Pro Pro
20      25      30
Phe Ile Glu Val Gly Lys Arg Lys Phe Arg Asp Trp Lys Lys Thr Gly
35      40      45
His Gly Asn Ser Asn Leu Tyr Lys Ala Ile Arg Glu Ser Val Asp Val
50      55      60
Tyr Phe Tyr Lys Phe Gly Leu Glu Ile Ser Ile Glu Xaa Leu Ser Lys
65      70      75      80
Xaa Phe Lys Xaa Ser Gly Leu Trp Gly Lys Asn Gly Arg
85      90

```

(2) INFORMATION FOR SEQ ID NO:149:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...60

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149

```

Met Ala His His Xaa Glu Gln His Gly Gly His His His His His
1      5      10      15
His Thr His His His His Tyr His Gly Gly Glu His His His His
20      25      30
His Ser Ser His His Glu Glu Gly Cys Cys Ser Thr Ser Asp Ser His
35      40      45
His Gln Glu Glu Gly Cys Cys His Gly Xaa His Glu
50      55      60

```

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 297 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

151

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...297

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:150

```

Leu Val Lys Ile Arg Leu Phe Asp Phe Thr Ile Arg Leu Phe Lys Pro
1      5      10      15
Glu Phe His Ile Phe Asp Phe Leu Lys Gly Ile Arg Val Leu Met Ile
      20      25      30
Glu Trp Met Gln Asn His Arg Lys Tyr Leu Val Val Thr Ile Trp Ile
      35      40      45
Ser Thr Ile Ala Phe Ile Ala Ala Gly Met Ile Gly Trp Gly Gln Tyr
      50      55      60
Ser Phe Ser Leu Asp Ser Asp Ser Ala Ala Lys Val Gly Gln Ile Lys
65      70      75      80
Ile Ser Gln Glu Glu Leu Ala Gln Glu Tyr Arg Arg Leu Lys Asp Ala
      85      90      95
Tyr Ala Glu Ser Ile Pro Asp Phe Lys Glu Leu Thr Glu Asp Gln Ile
      100      105      110
Lys Ala Met His Leu Glu Lys Ser Ala Leu Asp Ser Leu Ile Asn Gln
      115      120      125
Ala Leu Leu Arg Asn Phe Ala Leu Asp Leu Gly Leu Gly Ala Thr Lys
130      135      140
Gln Glu Val Ala Lys Glu Ile Arg Lys Thr Asn Val Phe Gln Lys Asp
145      150      155      160
Gly Val Phe Asp Glu Glu Leu Tyr Lys Asn Ile Leu Lys Gln Ser His
      165      170      175
Tyr Arg Pro Lys His Phe Glu Glu Ser Val Glu Arg Leu Leu Ile Leu
      180      185      190
Gln Lys Ile Ser Ala Leu Phe Pro Lys Thr Thr Thr Pro Leu Glu Gln
      195      200      205
Ser Ser Leu Ser Leu Trp Ala Lys Leu Gln Asp Lys Leu Asp Ile Leu
210      215      220
Ile Leu Asn Pro Asn Asp Val Lys Ile Ser Leu Asn Glu Glu Glu Met
225      230      235      240
Lys Lys Tyr Tyr Glu Asn His Arg Lys Asp Phe Lys Lys Pro Thr Ser
      245      250      255
Phe Lys Thr Arg Ser Leu Tyr Phe Asp Ala Ser Leu Glu Lys Thr Asp
      260      265      270
Leu Lys Glu Leu Glu Glu Tyr Tyr His Lys Asn Lys Val Ser Tyr Leu
275      280      285
Asp Xaa Xaa Gly Glu Ile Thr Gly Phe
290      295

```

## (2) INFORMATION FOR SEQ ID NO:151:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...90

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:151

```

Met Val Lys His Tyr Leu Phe Met Ala Val Ser Gln Val Phe Phe Ser
1      5      10      15
Phe Phe Leu Val Leu Phe Phe Ile Ser Ser Ile Val Leu Leu Ile Ser
      20      25      30

```

152

```

Ile Ala Ser Val Thr Leu Val Ile Lys Val Ser Phe Leu Asp Leu Val
      35              40              45
Gln Leu Phe Leu Tyr Ser Leu Pro Gly Thr Ile Phe Phe Ile Leu Pro
      50              55              60
Ile Thr Phe Phe Ala Ala Xaa Arg Leu Gly Xaa Ser Arg Leu Ser Tyr
65              70              75              80
Asp His Glu Leu Leu Val Phe Phe Leu Xaa
      85              90

```

## (2) INFORMATION FOR SEQ ID NO:152:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 86 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...86
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:152

```

Met Ser Lys Arg Ala Ile Arg Phe Pro Asn Lys Leu Phe Ser Tyr Pro
1              5              10              15
Lys Pro Lys Ile Lys Ala Thr Asn Thr Ser His Thr Val Leu Phe Ala
      20              25              30
Tyr Pro Leu Lys Pro His Glu Met Ala Leu Leu Ala Leu Ala Thr Ser
      35              40              45
Leu Leu Ala Pro Ile Phe Asn Ala Ile His Ser Thr Asn Ala Leu Asn
50              55              60
Ala Ile Lys Pro Asp Gly Thr Gly Ser Lys Ile Asn Pro Ile Ile Met
65              70              75              80
Pro Met Lys Ile Gln Lys
      85

```

## (2) INFORMATION FOR SEQ ID NO:153:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 141 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...141
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:153

```

Val Tyr Ser Arg Phe Phe Ala Asn Gln His Glu Phe Asp Phe Glu Ala
1              5              10              15
Gln Gly Ala Leu Gly Ser Asp Gln Ser Leu Asn Phe Lys Ser Thr
      20              25              30
Leu Leu Gln Asp Leu Asn Gln Ser Tyr Asn Tyr Leu Ala Tyr Ser Ala
      35              40              45
Thr Ala Arg Ala Ser Tyr Gly Tyr Asp Phe Ala Phe Phe Arg Asn Ala

```

153

```

      50              55              60
Leu Val Leu Lys Pro Ser Val Gly Val Ser Tyr Asn His Leu Gly Ser
65              70              75              80
Thr Asn Phe Lys Ser Asn Ser Gln Ser Gln Val Ala Leu Lys Asn Gly
      85              90              95
Ala Ser Ser Gln His Leu Phe Asn Ala Asn Ala Thr Trp Lys Arg Val
      100              105              110
Ile Ile Met Gly Thr Leu His Thr Phe Ile Cys Met Trp Glu Phe Tyr
      115              120              125
Lys Ser Ser Leu Thr Leu Asp Arg Met Met Trp Arg Leu
      130              135              140

```

## (2) INFORMATION FOR SEQ ID NO:154:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 185 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...185

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:154

```

Met Xaa Glu Asn Gly Arg Gly Val Pro Lys Asp Tyr Lys Lys Ala Val
1              5              10              15
Glu Tyr Phe Gln Lys Ala Val Asp Asn Asp Ile Pro Arg Gly Tyr Asn
      20              25              30
Asn Leu Gly Val Met Tyr Lys Glu Gly Lys Gly Val Pro Lys Asp Glu
      35              40              45
Lys Lys Ala Val Glu Tyr Phe Arg Ile Ala Thr Glu Lys Gly Tyr Thr
      50              55              60
Asn Ala Tyr Ile Asn Leu Gly Ile Met Tyr Met Glu Gly Arg Gly Val
65              70              75              80
Pro Ser Asn Tyr Ala Lys Ala Thr Glu Cys Phe Arg Lys Ala Met His
      85              90              95
Lys Gly Asn Val Xaa Ala Tyr Ile Leu Gly Asp Ile Tyr Tyr Ser
      100              105              110
Gly Met Ile Asn Trp Val Leu Ser Arg Thr Lys Ile Arg Leu Val His
      115              120              125
Tyr Lys Met Ala Ala Asp Val Ser Ser Ser Arg Ala Tyr Xaa Gly Leu
      130              135              140
Ser Glu Ser Tyr Xaa Tyr Gly Leu Gly Val Glu Lys Xaa Xaa Lys Lys
145              150              155              160
Ala Glu Glu Tyr Met Gln Lys Ala Cys Asp Phe Asp Ile Asp Lys Asn
      165              170              175
Cys Lys Lys Lys Asn Thr Ser Ser Arg
      180              185

```

## (2) INFORMATION FOR SEQ ID NO:155:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 139 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

154

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...139

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155

```

Leu  Leu  Asn  Met  Trp  Asp  Glu  Ala  Lys  Lys  Glu  Gly  Ile  Asn  Ile  Asn
1      5      10      15
Thr  Glu  Lys  Leu  Ser  Gln  Glu  Leu  Gly  Val  Val  Cys  Val  Pro  Thr  Ser
      20      25      30
Ala  Arg  Xaa  Lys  Glu  Asp  Arg  Leu  Asn  Thr  Glu  Leu  Leu  Asp  Glu
      35      40      45
Ile  Val  Arg  Leu  Tyr  Ser  Gln  Asn  Thr  Thr  Asn  Asn  Glu  Asn  Ile  Lys
      50      55      60
Val  Pro  Ser  Gln  Ser  Phe  Lys  Glu  Ser  Leu  Lys  Tyr  Ser  Gln  Ser  Ala
      65      70      75      80
Gln  Arg  Ile  Ala  Lys  Ser  Val  Ile  Ser  Glu  Asn  Lys  Gln  Asn  Ala  Ser
      85      90      95
Phe  Glu  His  Thr  Tyr  Lys  Ile  Asp  Lys  Ile  Phe  Asn  Ala  Pro  Ala  Leu
      100     105     110
Trp  Asp  Phe  His  Phe  Phe  Xaa  Val  Tyr  Val  Tyr  His  Leu  Phe  Phe  Glu
      115     120     125
Leu  Phe  Asn  Arg  Arg  Gly  Ser  Ala  Lys  Ser  Pro
      130     135

```

(2) INFORMATION FOR SEQ ID NO:156:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 193 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...193

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156

```

Met  Gln  Glu  Ala  Leu  Leu  Arg  Phe  Gln  Glu  Gly  Phe  Lys  Glu  Trp  Gly
1      5      10      15
Tyr  Leu  Ile  Leu  Phe  Leu  Tyr  Ser  Leu  Gly  Gly  Gly  Tyr  Val  Gly  Ile
      20      25      30
Val  Ile  Ala  Ser  Ile  Leu  Ser  Ala  Thr  Thr  His  Ala  Leu  Asp  Ile  Lys
      35      40      45
Ile  Thr  Ile  Leu  Val  Ala  Phe  Leu  Gly  Asn  Leu  Ile  Gly  Ser  Gly  Ala
      50      55      60
Leu  Val  Ile  Phe  Ala  Arg  Tyr  Gln  Lys  Arg  Glu  Phe  Leu  Lys  Tyr  Phe
      65      70      75      80
Gln  Lys  His  Arg  Arg  Lys  Leu  Ala  Leu  Ala  Ser  Leu  Trp  Val  Lys  Arg
      85      90      95
Tyr  Ala  Leu  Leu  Met  Ile  Phe  Val  Asn  Lys  Tyr  Leu  Tyr  Gly  Ile  Lys
      100     105     110
Ser  Val  Val  Pro  Leu  Ala  Ile  Gly  Phe  Ser  Lys  Tyr  Pro  Leu  Lys  Lys
      115     120     125
Phe  Leu  Trp  Leu  Asn  Val  Phe  Ser  Ser  Phe  Leu  Trp  Ala  Leu  Ile  Val
      130     135     140
Gly  Ser  Val  Ser  Phe  Gln  Ala  Ser  Asp  Trp  Val  Lys  Thr  Leu  Tyr  Glu
      145     150     155     160
Arg  Leu  Ser  His  Tyr  Thr  Ser  Phe  Phe  Val  Ile  Ser  Phe  Val  Leu  Ile
      165     170     175

```



155

Ala Leu Leu Ile Trp Phe Leu Leu Lys Arg Tyr Ser Arg Lys Met Gly  
                   180                  185                  190  
 Phe

## (2) INFORMATION FOR SEQ ID NO:157:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 129 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...129

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157

Met	Arg	Lys	Gly	Arg	Val	Met	Leu	Cys	Val	Phe	Asp	Ile	Glu	Thr	Ile
1				5					10				15		
Pro	Asn	Ile	Ser	Leu	Cys	Lys	Glu	His	Phe	Gln	Leu	Lys	Glu	Asp	Asp
			20				25						30		
Ala	Leu	Lys	Ile	Cys	Glu	Trp	Ser	Phe	Glu	Lys	Gln	Lys	Glu	Lys	Ser
		35				40					45				
Gly	Ser	Glu	Phe	Leu	Pro	Leu	Tyr	Leu	His	Glu	Ile	Ile	Ser	Ile	Ala
	50				55					60					
Ala	Val	Ile	Gly	Asp	Asp	Tyr	Gly	Gln	Phe	Ile	Lys	Val	Gly	Asn	Phe
65				70					75					80	
Gly	Gln	Lys	His	Glu	Asn	Lys	Glu	Asp	Phe	Ala	Ser	Glu	Lys	Glu	Leu
			85					90					95		
Leu	Glu	Asp	Phe	Phe	Lys	Tyr	Phe	Asn	Glu	Lys	Gln	Pro	Arg	Leu	Ile
		100					105					110			
Ser	Phe	Xaa	Gly	Arg	Gly	Phe	Gly	Tyr	Ser	Pro	Thr	His	Ala	Gln	Ser
		115					120						125		
Pro															

## (2) INFORMATION FOR SEQ ID NO:158:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 307 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...307

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158

Met	Ala	Lys	Lys	Lys	Ile	Ala	Ile	Ser	Cys	Gly	Asp	Ile	Gln	Gly	Val
1				5					10				15		
Gly	Leu	Glu	Leu	Ile	Leu	Lys	Ser	His	Lys	Glu	Val	Ser	Ala	Leu	Cys
		20					25					30			
Glu	Pro	Leu	Tyr	Leu	Val	His	Ser	Glu	Leu	Leu	Glu	Arg	Ala	Asn	Gln

SUBSTITUTE SHEET (RULE 26)

156

(2) INFORMATION FOR SEO ID NO:159:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 146 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Helicobacter pylori*

```
(ix) FEATURE:
      (A) NAME/KEY: misc_feature
      (B) LOCATION 1...146
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159

Met 1	Ser	Ser	Gly	Leu 5	Ile	Tyr	Ile	Ser	Leu 10	Glu	Val	Leu	Val	Xaa 15	Cys
Leu	Ile	Thr	Ala 20	Leu	Ile	Met	Tyr	Tyr 25	Val	Met	Lys	Lys	Ile 30	Tyr	Tyr
Ala	Arg	Gly 35	Gln	Ala	Ile	Leu	Lys 40	Gly	Ala	Ser	Ala	Lys 45	Ala	Lys	Leu
Met	Glu 50	Phe	Gln	Ala	Lys	Ser 55	Phe	Val	Glu	Ala	Glu 60	Glu	Met	Arg	Met
Lys 65	Ser	Gln	Glu	Cys	Lys 70	Leu	Gln	Gln	Gln	Tyr 75	Glu	Asn	Lys	Asn	Leu 80
Gln	Leu	Gln	Thr 85	His	Phe	Asp	Lys	Lys	Glu 90	Ala	His	Leu	Lys	His 95	Leu
Glu	Ala	Gln	His	Lys	Glu	Phe	Val	Arg	Asp	Glu	Lys	Arg	Tyr	Leu	Glu

157

```

          100          105          110
Lys Glu Lys Lys Glu Leu Glu Lys Glu Arg Gln Ile Leu Glu Xaa Glu
          115          120          125
Arg Glu Asn Phe Xaa Xaa Gln Arg Ala Phe Val Xaa Xaa Xaa Xaa Ala
          130          135          140
Lys Ala
145

```

## (2) INFORMATION FOR SEQ ID NO:160:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 127 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...127

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160

```

Met Asn Ile Lys Ile Leu Lys Ile Leu Val Gly Gly Leu Phe Phe Leu
1      5      10
Ser Leu Asn Ala His Leu Trp Gly Lys Gln Asp Asn Ser Phe Leu Gly
20     25     30
Ile Gly Glu Arg Ala Tyr Lys Ser Gly Asn Tyr Ser Lys Ala Ala Ser
35     40     45
Tyr Phe Lys Lys Ala Cys Asn Asp Gly Val Ser Glu Gly Cys Thr Gln
50     55     60
Leu Gly Ile Ile Tyr Glu Asn Gly Gln Gly Thr Arg Ile Asp Tyr Lys
65     70     75     80
Lys Ala Leu Glu Tyr Tyr Lys Thr Ala Cys Gln Ala Asp Asp Arg Glu
85     90     95
Gly Cys Phe Gly Leu Gly Gly Leu Tyr Asp Glu Gly Leu Gly Thr Ala
100    105    110
Gln Asn Tyr Gln Glu Ala Xaa Asp Ala Tyr Ala Arg His Ala Phe
115    120    125

```

## (2) INFORMATION FOR SEQ ID NO:161:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 116 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...116

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161

```

Met Ala Ile Ala Ile Lys Asp Leu Leu Ser Ala Tyr Lys Val Val Leu
1      5      10      15
Pro Leu Asp Lys Ile Ser Met Pro Ser Ser Ala Asp Leu Lys Leu Thr
20     25     30

```

158

```

Leu Gln Phe Leu Lys Asn Thr Ala Pro Leu Phe Ser Val Gln Gly Ser
      35              40              45
Val Asn Leu Gln Glu Gly Thr Phe Ser Leu Tyr Asn Ile Pro Leu Tyr
      50              55              60
Thr Gln Ser Ala Gln Ile Asn Leu Asp Ile Ala Gln Glu Tyr Gln Tyr
      65              70              75              80
Ile Tyr Ile Asp Thr Ile His Thr Arg Tyr Ala Asn Met Xaa Asp Leu
      85              90              95
Asp Ala Lys Ile Ala Leu Asp Leu Gly Gln Lys Asn Leu Ser Xaa Xaa
      100             105             110
Xaa Leu Gly Pro
      115

```

## (2) INFORMATION FOR SEQ ID NO:162:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 82 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...82
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:162

```

Met Pro Asp Asn Leu His Leu His Thr Leu Leu Xaa Lys Phe Leu Gln
1      5      10
Gln Arg Ser Phe Asn Tyr Pro Asn Pro Leu Cys Ala Phe Ile Leu Ile
      20      25      30
Leu Cys Asn Leu Pro Phe Ile Leu Ile Ser Val Leu Phe Arg Leu Asp
      35      40      45
Ala Tyr Ala Leu Ile Val Ile Ser Leu Val Phe Ile Xaa Cys Tyr Leu
      50      55      60
Ile Gly Xaa Ala Tyr Leu Asn Arg Gln Val Cys Ala Leu Glu Lys Arg
      65      70      75      80
Ala Phe

```

## (2) INFORMATION FOR SEQ ID NO:163:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 116 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...116
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:163

```

Leu Lys Val Thr Asn Pro His Leu Leu Val Val Ile Gln Asp Leu Asn
1      5      10      15
Ala Arg Ile Ala Leu Met Lys Leu Leu Phe Gln Asn Val Lys Ser Ala

```

[illegible]

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: YES

```
(ix) FEATURE:
      (A) NAME/KEY: misc_feature
      (B) LOCATION 1...198
```

[illegible]

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 85 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

**SUBSTITUTE SHEET (RULE 26)**

160

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...85

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165

```

Val Gly Ile Val Pro Asp Asn Leu Trp Lys Leu Lys Arg Phe Asn Gln
1      5      10      15
Asp Trp Arg Val Gly Asp Thr Leu Ile Thr Ala Ile Gly Gln Gly Ser
20      25      30
Phe Leu Ala Thr Pro Leu Gln Val Leu Ala Tyr Thr Gly Leu Ile Ala
35      40      45
Thr Gly Lys Leu Ala Thr Pro His Phe Ala Ile His Asn Gln Gln Pro
50      55      60
Leu Lys Asp Pro Leu Asn Arg Phe Ser Lys Lys Glu Ala Pro Ser Leu
65      70      75      80
Ala Arg Gly His Val
85

```

(2) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 343 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...343

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166

```

Met Gln Asn Leu Leu Ile Gln Ala Glu Asn Ala Ile Ala Leu Leu Phe
1      5      10      15
Leu Leu Asn Asp Lys Asn Leu Lys Gly Lys Ile Asp Leu Ile Tyr Ile
20      25      30
Asp Pro Pro Phe Ala Thr Asn Asn His Phe Thr Ile Thr Asn Gly Arg
35      40      45
Ala Thr Thr Ile Ser Asn Ser Lys Asn Gly Asp Ile Ala Tyr Ser Asp
50      55      60
Lys Val Val Gly Met Asp Phe Met Glu Phe Leu Lys Gln Arg Leu Val
65      70      75      80
Leu Leu Lys Glu Leu Leu Ser Glu Gln Gly Ser Ile Tyr Val His Thr
85      90      95
Asp Tyr Lys Ile Gly His Tyr Val Lys Val Met Leu Asp Glu Ile Phe
100     105     110
Gly Ile Gln Asn Phe Arg Asn Glu Ile Thr Arg Ile Lys Cys Asn Pro
115     120     125
Lys Asn Phe Lys Arg Ile Gly Tyr Gly Asn Ile Lys Asp Met Ile Leu
130     135     140
Phe Tyr Ser Lys Gly Lys Asn Pro Ile Phe Asn Glu Pro Lys Ile Pro
145     150     155     160
Tyr Thr Pro Gln Asp Leu Glu Lys Arg Phe Pro Lys Ile Asp Lys Asp
165     170     175
Lys Arg Arg Tyr Thr Thr Val Pro Ile His Ala Pro Gly Glu Val Glu
180     185     190

```

161

```

Ser Gly Glu Cys Ser Lys Ala Phe Lys Gly Met Leu Pro Pro Lys Gly
    195      200      205
Arg His Trp Arg Thr Asp Ile Ala Thr Leu Glu Arg Trp Asp Lys Glu
    210      215      220
Gly Leu Ile Glu Tyr Ser Asn Asn Asn Pro Arg Lys Lys Ile Tyr
    225      230      235      240
Ala Leu Glu Gln Val Gly Lys Arg Val Gln Asp Ile Trp Glu Phe Lys
    245      250      255
Asp Pro Gln Tyr Pro Ser Tyr Pro Thr Glu Lys Asn Ala Gln Leu Leu
    260      265      270
Asp Leu Ile Ile Lys Thr Ser Ser Asn Lys Asp Ser Ile Val Leu Asp
    275      280      285
Cys Phe Cys Gly Ser Gly Thr Thr Leu Lys Ser Ala Phe Leu Leu Gln
    290      295      300
Arg Lys Phe Ile Gly Ile Asp Asn Ser Asp Leu Ala Ile Gln Ala Cys
    305      310      315      320
Lys Asn Lys Leu Glu Thr Ile Thr Lys Asp Leu Phe Val Ser Gln Asn
    325      330      335
Phe Tyr Asp Phe Leu Val Phe
    340

```

## (2) INFORMATION FOR SEQ ID NO:167:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 176 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...176

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:167

```

Met Thr Ser Val Val Ile Lys Pro His Ala Tyr Gly Glu Gln Val Gln
 1      5      10      15
Glu Ile Glu Glu Ser Asp Ser Asp Tyr Glu Lys Asn Asn Asp Gln
    20      25      30
Glu Ala Ile Asn Phe Gly Ile Ala Leu His Lys Gly Leu Glu Tyr Gln
    35      40      45
Tyr Ala Tyr Asn Ile Pro Lys Gln Ser Val Leu Glu Tyr Leu Asn Tyr
    50      55      60
His Tyr Gly Phe Tyr Gly Leu Asp Tyr Gln Ala Leu Glu Glu Ser Leu
    65      70      75      80
Glu Leu Phe Glu Asn Asp Ala Gly Ile Gln Ala Leu Phe Lys Asn His
    85      90      95
Ala Leu Lys Gly Glu Ala Ala Phe Leu Phe Gln Gly Val Val Ser Arg
    100      105      110
Ile Asp Val Leu Leu Trp Asp Arg Gly Gln Asn Leu Tyr Val Leu Asp
    115      120      125
Tyr Lys Ser Ser Gln Asn Tyr Gln Gln Ser His Lys Ala Gln Val Ser
    130      135      140
His Tyr Ala Glu Phe Leu Arg Thr Gln Xaa Pro His Phe Lys Ile Gln
    145      150      155      160
Ala Gly Ile Ile Tyr Ala His Lys Arg Leu Leu Glu Lys Xaa Trp Xaa
    165      170      175

```

## (2) INFORMATION FOR SEQ ID NO:168:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 260 amino acids
- (B) TYPE: amino acid

SUBSTITUTE SHEET (RULE 26)

162

- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...260
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:168

Met	Ser	Glu	Asp	Leu	Pro	Phe	Ala	Ser	Asp	Ser	Gln	Phe	Thr	Tyr	Asn
1				5					10					15	
Gly	Val	Ser	Ile	Thr	Arg	Pro	Thr	Asn	Glu	Val	Asn	Asp	Val	Ile	Ser
			20					25					30		
Gly	Val	Asn	Ile	Thr	Leu	Glu	Gln	Thr	Thr	Glu	Pro	Asn	Lys	Pro	Ala
		35					40					45			
Ile	Ile	Ser	Val	Ser	Arg	Asp	Asn	Gln	Ala	Ile	Ile	Asp	Ser	Leu	Lys
	50					55					60				
Glu	Phe	Val	Lys	Ala	Tyr	Asn	Glu	Leu	Ile	Pro	Lys	Leu	Asp	Glu	Asp
65					70					75				80	
Thr	Arg	Tyr	Asp	Ala	Asp	Thr	Lys	Ile	Ala	Gly	Ile	Phe	Asn	Gly	Val
			85						90				95		
Gly	Asp	Ile	Arg	Ala	Ile	Arg	Ser	Ser	Leu	Asn	Asn	Val	Phe	Ser	Tyr
			100					105					110		
Ser	Val	His	Thr	Asp	Asn	Gly	Val	Glu	Ser	Leu	Met	Lys	Tyr	Gly	Leu
		115				120						125			
Ser	Leu	Asp	Asp	Lys	Gly	Val	Met	Ser	Leu	Asp	Glu	Ala	Lys	Leu	Ser
	130					135					140				
Ser	Ala	Leu	Asn	Ser	Asn	Pro	Lys	Ala	Thr	Gln	Asp	Phe	Phe	Tyr	Gly
145					150					155					160
Ser	Asp	Ser	Lys	Asp	Met	Gly	Gly	Arg	Glu	Ile	His	Gln	Glu	Gly	Ile
				165					170					175	
Phe	Ser	Lys	Phe	Asn	Gln	Val	Ile	Ala	Asn	Leu	Ile	Asp	Gly	Gly	Asn
			180					185					190		
Ala	Lys	Leu	Lys	Ile	Tyr	Glu	Asp	Ser	Leu	Asp	Arg	Asp	Ala	Lys	Ser
		195					200					205			
Leu	Thr	Lys	Asp	Lys	Glu	Asn	Ala	Gln	Glu	Leu	Leu	Lys	Thr	Arg	Tyr
	210					215					220				
Asn	Ile	Met	Ala	Glu	Arg	Phe	Ala	Ala	Tyr	Asp	Ser	Gln	Ile	Ser	Lys
225					230					235				240	
Ala	Asn	Gln	Lys	Phe	Asn	Ser	Val	Gln	Met	Met	Ile	Asp	Gln	Ala	Ala
			245					250						255	
Ala	Lys	Lys	Asn												
			260												

## (2) INFORMATION FOR SEQ ID NO:169:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 145 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...145
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:169



163

```

Met Arg Ile Val Phe Met Gly Thr Pro Ser Phe Ala Glu Val Ile Leu
1      5      10      15
Arg Ala Leu Val Glu Asn Glu Asp Lys Lys Ile Glu Val Val Gly Leu
      20      25      30
Phe Thr Gln Arg Asp Lys Pro Phe Gly Arg Lys Lys Glu Leu Lys Ala
      35      40      45
Pro Glu Thr Lys Thr Tyr Ile Leu Glu Asn His Leu Asn Ile Pro Ile
      50      55      60
Phe Gln Pro Gln Ser Leu Lys Glu Pro Glu Val Gln Ile Leu Lys Gly
65      70      75      80
Leu Lys Pro Asp Phe Ile Val Val Val Ala Tyr Gly Lys Ile Leu Pro
      85      90      95
Lys Glu Val Leu Thr Ile Ala Pro Cys Ile Asn Leu His Ala Ser Leu
      100      105      110
Leu Pro Lys Tyr Arg Gly Ala Ser Pro Ile His Glu Met Ile Leu Asn
      115      120      125
Asp Asp Arg Ile Tyr Gly Ile Ser Thr Met Leu Met Xaa Phe Gly Ile
130      135      140
Gly
145

```

## (2) INFORMATION FOR SEQ ID NO:170:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 248 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...248

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:170

```

Met Arg Phe Tyr Phe Lys Phe Leu Trp Leu Leu Gly Ile Phe Leu Ile
1      5      10      15
Phe Tyr Phe Leu Asp Ile Lys Gly Ser Ser Ser Tyr Ile Ser Asp Arg
      20      25      30
Val Lys Ser Ala Leu Met Ser Ala Lys Asn Ser Leu Leu Asp Asn Val
      35      40      45
Gln Ala Tyr Phe Phe Gln Ala Gln Asn Ile Lys Glu Phe Gln Lys Glu
      50      55      60
Arg Leu Ile Leu Glu Ala Leu Lys Leu Glu Asn Ala Asp Leu Lys Glu
65      70      75      80
Arg Leu Asn Ser Ile Tyr Pro Leu Glu Asn Pro Lys Met Thr Tyr Thr
      85      90      95
Pro Thr Phe Met Thr Ser Phe Ile Asn Leu Glu Asp Thr His Ser Val
      100      105      110
Ser Leu Asn Pro Ile Val Asn Leu Glu Glu Asn Lys Ile Tyr Gly Leu
      115      120      125
Val Ser His Asn Gln Ala Ile Gly Ile Ala Val Leu Glu Lys Gly Arg
130      135      140
Leu Asn Gly Phe Leu Asn Ala His Lys Arg Cys Ala Tyr Ser Val Met
145      150      155      160
Ile Gly Gln Asn Gln Val Leu Gly Phe Ile Gly Thr Asn Phe Lys Gln
      165      170      175
Glu Leu Val Val Asp Phe Ile Val Pro Ser Ala Glu Ile Asn Ile Gly
      180      185      190
Asp Gln Val Leu Thr Ser Gly Leu Asp Gly Ile Phe Gly Ala Gly Val
195      200      205
Phe Val Gly Glu Val Ser Ser Val Glu Asp His Tyr Thr Tyr Lys Ser

```

SUBSTITUTE SHEET (RULE 26)

164

```

      210              215              220
Ala Val Leu Lys Asn Ala Phe Leu Ser Glu Ala Lys Leu Leu Arg His
225              230              235              240
Val Phe Leu Ser Gly Val Lys Asn
      245

```

## (2) INFORMATION FOR SEQ ID NO:171:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171

```

Leu Ala Leu Arg Leu Pro Phe Leu Ile Ala His Val Ile Asn Met Phe
1              5              10              15
Leu Phe Tyr Leu Ile Gly Arg Lys Ile Leu Lys Lys Pro Lys Asp Ala
20              25              30
Leu Tyr Val Val Leu Thr Tyr Ala Leu Leu Pro Gly Val Asn Leu Phe
35              40              45
Ala Ile Leu Leu Ala Lys Ser Val Leu Val Leu Ser Leu Gly Leu Leu
50              55              60
Ile Ser Tyr Leu Tyr Ile Lys Thr Gln Lys Ile Pro Tyr Leu Thr Leu
65              70              75              80
Ser Ala Cys Ala Phe Leu Asp Gly Ala Phe Ile Pro Leu Leu Leu Gly
85              90              95
Val Phe Ala Tyr Ala Leu Arg Lys Thr Ala Ile Leu Arg Ala Arg Ser
100              105              110
Leu Leu Trp Trp Phe Xaa Leu
115

```

## (2) INFORMATION FOR SEQ ID NO:172:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 108 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...108

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172

```

Val Asn Leu Met Asp Tyr Phe Ser Lys Ser Leu Phe Leu Asn Ser Leu
1              5              10              15
Asn Thr Gln Arg Leu Ile Val Ser Asn Lys Leu Ala Ile Asp Val Gln
20              25              30
Tyr Gly Met Leu Gln Ser Val Arg Lys Asn Tyr Pro Asp Val Val Asp
35              40              45

```

165

```

Gly Gly Val Arg Glu Gly Pro Phe Trp Val Leu Ala Gly Ala Leu Met
 50      55      60
Pro Ser Ile Leu Ile Glu Ile Gly Tyr Asn Ser His Ala Ile Glu Ser
65      70      75      80
Lys Arg Ile Gln Ser Lys Pro Tyr Gln Lys Ile Leu Ala Lys Gly Ile
      85      90      95
Ala Asp Gly Ile Asp Ser Phe Phe Ser Lys Asn Asp
      100      105

```

## (2) INFORMATION FOR SEQ ID NO:173:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 157 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...157

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:173

```

Leu Ala Ser Arg Tyr Ser Val Ala Val Gly Asn Leu Phe Ser Glu His
 1      5      10      15
Leu Tyr Asp Leu Arg Asn Glu Thr Met Thr Asn Leu Ile Gly Phe Leu
      20      25      30
Leu Val Leu Ala Ser Ile Trp Val Phe Phe Leu Ala Leu Gly Val Leu
      35      40      45
Leu Gly Lys Met Leu Val Phe Ser Gly Leu Gly Ile Ile Asp Lys Ala
      50      55      60
Leu Gly Phe Ile Phe Ser Cys Leu Lys Thr Phe Leu Val Leu Ser Phe
65      70      75      80
Ile Leu Tyr Ala Leu Ser Lys Met Asp Leu Met Lys Asp Ala Asn Ala
      85      90      95
Tyr Leu Gln Glu Lys Xaa Xaa Ile Phe Pro Thr Xaa Lys Xaa Xaa Xaa
      100      105      110
Ser Lys Ile Met Arg Leu Asp Gly Val Lys His Val Glu Lys Asn Leu
      115      120      125
Lys Asp Asn Leu Glu Glu Met Ser Asp Glu Val Lys Asn Lys Gly Ser
      130      135      140
Ile Asp Asn Ala Lys Glu Ser Phe Asn Lys Gly Tyr Gly
145      150      155

```

## (2) INFORMATION FOR SEQ ID NO:174:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...81

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:174

166

```

Leu Ser Lys Gln Ser Ala Asp Ile Val Ile Thr Asn Asp Ser Leu Ser
1          5          10          15
Ser Leu Val Lys Val Leu Ala Ile Ala Lys Lys Thr Lys Ser Ile Thr
20          25          30
Trp Gln Asn Ile Leu Phe Ala Leu Gly Ile Lys Ala Val Phe Ile Val
35          40          45
Leu Gly Leu Met Gly Val Ala Ser Leu Trp Glu Ala Val Phe Gly Asp
50          55          60
Val Gly Val Thr Leu Leu Ala Leu Ala Asn Ser Xaa Arg Thr Met Arg
65          70          75          80
Ala

```

## (2) INFORMATION FOR SEQ ID NO:175:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 80 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...80

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:175

```

Met Lys Asn Leu Arg His Phe Arg Lys Leu Ile Ala Phe Leu Gly Phe
1          5          10          15
Ser Pro Leu Leu Gln Ala Asp Met Thr Thr Phe Phe Asn Ser Ile
20          25          30
Glu Gln Gln Leu Thr Ser Pro Thr Ala Lys Gly Ile Leu Met Val Ile
35          40          45
Phe Leu Gly Leu Ala Ile Phe Ile Trp Lys Asn Leu Asp Arg Trp Lys
50          55          60
Glu Ile Leu Met Thr Val Leu Ala Leu Lys Xaa Val Pro Met Gln Xaa
65          70          75          80

```

## (2) INFORMATION FOR SEQ ID NO:176:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 325 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...325

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:176

```

Leu Ala Gly Leu Xaa Val Gly Cys Xaa Arg Met Lys Gln Thr Phe Trp
1          5          10          15
Xaa Leu Ser Trp Gly Glu Lys Ser Gln Lys Val Cys Val His Arg Pro
20          25          30

```

167

```

Trp Tyr Ala Ile Trp Ser Cys Asp Lys Trp Glu Glu Lys Thr Gln Gln
   35           40           45
Phe Thr Gly Asn Gln Leu Ile Thr Lys Thr Trp Ala Gly Gly Asn Ala
   50           55           60
Ala Asn Tyr Tyr His Ser Gln Asn Asn Gln Asp Ile Thr Ala Asn Leu
   65           70           75           80
Lys Asn Asp Asn Gly Thr Tyr Phe Leu Ser Gly Leu Tyr Asn Tyr Thr
           85           90           95
Gly Gly Glu Tyr Asn Gly Gly Asn Leu Asp Ile Glu Leu Gly Ser Asn
           100          105          110
Ala Thr Phe Asn Leu Gly Ala Ser Ser Gly Asn Ser Phe Thr Ser Trp
           115          120          125
Tyr Pro Asn Gly His Thr Asp Val Thr Phe Ser Ala Gly Thr Ile Asn
           130          135          140
Val Asn Asn Ser Val Glu Val Gly Asn Arg Val Gly Ser Gly Ala Gly
           145          150          155          160
Thr His Thr Gly Thr Ala Thr Leu Asn Leu Asn Ala Asn Lys Val Thr
           165          170          175
Ile Asn Ser Asn Ile Ser Ala Tyr Lys Thr Ser Gln Val Asn Val Gly
           180          185          190
Asn Ala Asn Ser Val Ile Thr Ile Asn Ser Val Ser Leu Asn Gly Glu
           195          200          205
Thr Cys Xaa Ser Leu Ala Arg Val Gly Val Gly Ala Asn Cys Ser Thr
           210          215          220
Ser Gly Pro Ser Tyr Ser Phe Lys Gly Thr Thr Asn Ala Thr Asn Thr
           225          230          235          240
Thr Phe Ser Xaa Ser Ser Gly Ser Phe Thr Phe Glu Glu Asn Ala Thr
           245          250          255
Phe Ser Gly Ala Lys Leu Asn Gly Gly Ala Phe Thr Phe Asn Lys Lys
           260          265          270
Phe Asn Ala Thr Asn Asn Thr Ala Phe Asn Ser Gly Ser Phe Thr Phe
           275          280          285
Lys Gly Thr Ser Ser Phe Asn Gly Ala Asn Phe Ser Asn Ala Ser Tyr
           290          295          300
Thr Phe Asn Asn Gln Ala Thr Phe Gln Asn Ser Ser Phe Asn Gly Gly
           305          310          315          320
Thr Phe Thr Phe Asn
           325

```

## (2) INFORMATION FOR SEQ ID NO:177:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 271 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...271

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:177

```

Leu Leu Ser Leu Val Lys Gly Lys Thr Met Leu Arg Ser Leu Tyr Ser
1           5           10           15
Ala Thr Ser Gly Met Leu Ala Gln Gln Thr His Ile Asp Thr Thr Ser
           20           25           30
Asn Asn Ile Ala Asn Val Asn Thr Thr Gly Phe Lys Lys Ser Arg Ala
           35           40           45
Asp Phe Asn Asp Leu Phe Tyr Gln Ala Met Gln Tyr Ala Gly Thr Asn
           50           55           60
Thr Ser Asn Thr Thr Leu Ser Pro Asp Gly Met Glu Val Gly Leu Gly
           65           70           75           80

```

SUBSTITUTE SHEET (RULE 26)

168

```

Val Arg Pro Ser Ala Ile Thr Lys Met Phe Ser Gln Gly Ser Pro Lys
      85      90      95
Glu Thr Glu Asn Asn Leu Asp Ile Ala Ile Thr Gly Lys Gly Phe Phe
      100      105      110
Gln Val Gln Leu Pro Asp Gly Thr Thr Ala Tyr Thr Arg Ser Gly Asn
      115      120      125
Phe Lys Leu Asp Glu Gln Gly Asn Leu Val Thr Ser Glu Gly Tyr Leu
      130      135      140
Leu Ile Pro Gln Ile Thr Leu Pro Glu Asp Thr Thr Gln Val Asn Ile
      145      150      155      160
Gly Val Asp Gly Thr Val Ser Val Thr Gln Gly Leu Gln Thr Thr Ser
      165      170      175
Asn Val Ile Gly Gln Ile Thr Leu Ala Asn Phe Val Asn Pro Ala Gly
      180      185      190
Leu His Ser Met Gly Asp Asn Leu Phe Ser Ile Thr Asn Ala Ser Gly
      195      200      205
Asp Ala Ile Val Gly Asn Pro Asp Ser Gln Gly Leu Gly Lys Leu Arg
      210      215      220
Gln Gly Phe Leu Glu Leu Ser Asn Val Arg Leu Val Glu Glu Met Thr
      225      230      235      240
Asp Leu Ile Thr Ala Gln Arg Ala Tyr Glu Ala Asn Ser Lys Ser Ile
      245      250      255
Gln Thr Ala Asp Ala Met Leu Gln Thr Val Asn Ser Leu Lys Arg
      260      265      270

```

## (2) INFORMATION FOR SEQ ID NO:178:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...90

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178

```

Met Gln Asn Gly Tyr Tyr Ala Ala Thr Gly Ala Met Ala Thr Gln Phe
1      5      10
Asn Arg Leu Asp Leu Thr Ser Asn Asn Leu Ala Asn Leu Asn Thr Asn
      20      25      30
Gly Phe Lys Arg Asp Asp Ala Ile Thr Gly Asp Phe Leu Arg Leu Tyr
      35      40      45
Gln Glu Tyr Arg Glu Gln Leu Pro Leu Glu Asp Gln Thr Lys Ala Ser
      50      55      60
Ala Lys Tyr Leu Asn Arg Xaa Leu Asn Arg Val Pro Ile Leu Ser Xaa
65      70      75      80
Ile Tyr Thr Xaa Arg Xaa Leu Gly Xaa Val
      85      90

```

## (2) INFORMATION FOR SEQ ID NO:179:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

169

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...195

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179

```

Val Gly Ala Met Pro Thr Ile Gln Ile Arg Xaa Phe Gly Ala Gly Gly
1      5      10      15
Ser Gly His Ser Asp Ala Thr Leu Met Leu Val Asn Gly Ile Pro Val
      20      25      30
Tyr Met Ala Pro Tyr Ala His Ile Glu Leu Asp Ile Phe Pro Val Thr
      35      40      45
Phe Gln Ala Ile Asp Arg Ile Asp Val Ile Lys Gly Gly Gly Ser Val
50      55      60
Gln Tyr Gly Pro Asn Thr Tyr Gly Gly Ile Val Asn Ile Ile Thr Lys
65      70      75      80
Pro Ile Pro Asn Gln Trp Glu Asn Gln Ala Ala Glu Arg Xaa Thr Tyr
      85      90      95
Trp Ala Lys Ala Arg Asn Ala Gly Phe Ala Ala Pro Xaa Asp Lys Thr
      100      105      110
Gly Asp Pro Ser Phe Ile Lys Ser Leu Gly Asn Asn Leu Leu Tyr Asn
      115      120      125
Thr Tyr Val Arg Ser Gly Gly Met Ile Asn Lys His Val Gly Ile Gln
130      135      140
Arg Lys Leu Thr Gly Leu Glu Ala Lys Ala Leu Gly Thr Ile Ala Pro
145      150      155      160
Leu Val Phe Gln Thr Ile Gly Trp Met Gly Ser Met Thr Ser Met Lys
      165      170      175
Ala Met Gly Leu Lys Pro Ile Thr Asn Thr Thr Ile Leu Ala Ile Xaa
180      185      190
Gln Pro Gly
195
  
```

(2) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 84 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...84

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180

```

Met Arg Xaa Glu Lys Ile Met Thr Asn Phe Glu Lys Xaa Ile Ala Gln
1      5      10      15
Asn Arg Leu Lys Thr Asn Ala Val Leu Thr Thr Tyr Cys Ala Ile Phe
      20      25      30
Ala Phe Ile Gly Leu Leu Val Asp Ala Ile Arg Ile Asn Ala Asn Asp
      35      40      45
Leu Gly Ile Ala Leu Phe Lys Leu Met Thr Phe Gln Ile Phe Pro Thr
50      55      60
Xaa Thr-Ile Val Met Phe Val Val Ala Phe Val Ile Xaa Xaa Ser Leu
65      70      75      80
Tyr Pro Lys Phe
  
```

170

## (2) INFORMATION FOR SEQ ID NO:181:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 76 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...76

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:181

```

Met Xaa Met Ser His Ile Ile Lys Ser Ile Glu Ala Leu Asp Asp Tyr
1      5      10
Thr Ile Arg Phe Thr Leu Asn Gly Pro Glu Ala Pro Phe Leu Ala Asn
20      25      30
Leu Gly Met Asp Phe Leu Ser Ile Leu Ser Lys Asp Tyr Ala Asp Tyr
35      40      45
Leu Ala Gln Asn Asn Lys Lys Asp Glu Leu Ala Lys Xaa Pro Val Gly
50      55      60
Thr Gly Pro Phe Lys Phe Phe Leu Trp Asn Lys Arg
65      70      75

```

## (2) INFORMATION FOR SEQ ID NO:182:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 196 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...196

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:182

```

Leu Met Arg Lys Ile Phe Ser Tyr Ile Ser Lys Val Leu Leu Phe Ile
1      5      10      15
Gly Val Val Tyr Ala Glu Pro Asp Ser Lys Val Glu Ala Leu Glu Gly
20      25      30
Arg Lys Gln Glu Ser Ser Leu Asp Lys Lys Ile Arg Gln Glu Leu Lys
35      40      45
Ser Lys Glu Leu Lys Asn Lys Glu Leu Lys Asn Lys Asp Leu Lys Asn
50      55      60
Lys Glu Glu Lys Lys Glu Thr Lys Ala Lys Arg Lys Pro Arg Ala Glu
65      70      75      80
Val His His Gly Asp Ala Lys Asn Pro Thr Pro Lys Ile Thr Pro Pro
85      90      95
Lys Ile Lys Gly Ser Ser Lys Gly Val Gln Asn Gln Gly Val Gln Asn
100      105      110
Asn Ala Pro Lys Pro Glu Glu Lys Asp Thr Thr Pro Gln Ala Thr Glu
115      120      125
Lys Asn Lys Glu Thr Ser Pro Ser Ser Gln Phe Asn Ser Ile Phe Gly
130      135      140

```



171

```

Asn Pro Asn Asn Ala Thr Asn Asn Thr Leu Glu Asp Lys Val Val Gly
145          150          155          160
Gly Ile Ser Leu Leu Val Asn Gly Ser Pro Ile Thr Leu Tyr Gln Ile
          165          170          175
Gln Glu Glu Gln Glu Lys Ser Lys Val Ser Xaa Ala Xaa Ala Arg Asp
          180          185          190
Arg Leu Xaa Xaa
          195

```

## (2) INFORMATION FOR SEQ ID NO:183:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 179 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...179

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:183

```

Met Lys Arg Ser Ser Val Phe Ser Phe Leu Val Ala Phe Leu Leu Val
1      5      10      15
Val Gly Cys Ser His Lys Met Asp Asn Lys Thr Val Ala Gly Asp Val
20      25      30
Ser Thr Lys Ala Val Gln Thr Ala Pro Val Thr Thr Glu Pro Ala Pro
35      40      45
Glu Lys Glu Glu Pro Lys Gln Glu Pro Ala Pro Val Val Glu Glu Lys
50      55      60
Pro Ala Ile Glu Ser Gly Thr Ile Ile Ala Ser Ile Tyr Phe Asp Phe
65      70      75      80
Asp Lys Tyr Glu Ile Lys Glu Ser Asp Gln Glu Thr Leu Asp Glu Ile
85      90      95
Val Gln Lys Ala Lys Glu Asn His Met Gln Val Leu Leu Glu Gly Asn
100     105     110
Thr Asp Glu Phe Gly Ser Ser Glu Tyr Asn Gln Ala Leu Gly Val Lys
115     120     125
Arg Thr Leu Ser Val Lys Asn Ala Leu Val Ile Lys Gly Val Glu Lys
130     135     140
Asp Met Ile Lys Thr Ile Ser Phe Gly Glu Ser Lys Pro Lys Cys Val
145     150     155     160
Gln Lys Thr Arg Glu Cys Tyr Arg Glu Asn Arg Arg Val Asp Val Lys
165     170     175
Leu Val Lys

```

## (2) INFORMATION FOR SEQ ID NO:184:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 286 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

172

(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...286

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:184

```

Met Gly Thr Leu Ile Glu Lys Trp Phe Gly Phe Ser Gln Ile Arg Glu
1      5      10      15
Glu Leu Glu Ala Arg Ile Ser Glu Leu Glu Asp Glu Asn Thr Glu Leu
20      25      30
Leu Arg Glu Arg Glu Tyr Leu Ala Ala Glu Thr Ser Glu Leu Lys Asp
35      40      45
Ala Asn Asp Gln Leu Arg Gln Lys Asn Asp Lys Leu Phe Ile Thr Lys
50      55      60
Asp Lys Leu Thr Lys Glu Asn Thr Glu Leu Phe Ala Glu Asn Glu Ser
65      70      75      80
Leu Ser Val Lys Ile Ser Gly Leu Glu His Ser Asn Asp Gln Leu Trp
85      90      95
Gln Asn Asn Asn Lys Leu Thr Lys Glu Lys Ala Glu Leu Lys Thr Glu
100     105     110
Lys Asp Ile Leu Ala Lys Glu Asn Thr Arg Leu Leu Ala Ala Arg Asp
115     120     125
Arg Leu Thr Glu Glu Lys Arg Glu Leu Thr Thr Glu Lys Glu Arg Leu
130     135     140
Lys Arg Glu Asn Thr Glu Leu Thr His Lys Ile Thr Glu Leu Thr Lys
145     150     155     160
Glu Asn Lys Ala Leu Thr Thr Glu Asn Asp Lys Leu Asn His Gln Val
165     170     175
Thr Ala Leu Thr Asn Glu Arg Asp Ser Leu Glu Gln Glu Arg Ala Arg
180     185     190
Leu Gln Asp Ala His Gly Phe Leu Glu Lys Arg Cys Thr Asn Leu Glu
195     200     205
Lys Glu Asn Gln Arg Leu Thr Asp Lys Leu Lys Gln Leu Glu Ser Ala
210     215     220
Gln Lys Ser Leu Glu Asn Thr Asn Asn Gln Leu Arg Gln Ala Leu Glu
225     230     235     240
Asn Ser Asn Val Gln Leu Ala Gln Ala Lys Glu Xaa Ile Ala Ile Glu
245     250     255
Xaa Ser Glu Leu Xaa Arg Arg Asn Arg Thr Leu Glu Glu Leu Arg Gly
260     265     270
Tyr Gly Ser Gln Lys Xaa Ile Trp Thr Tyr Thr Xaa Gly Val
275     280     285

```

## (2) INFORMATION FOR SEQ ID NO:185:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...110

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:185

```

Val Leu Arg Lys Leu Leu Gly Lys Asn Cys Ile Glu Thr His Lys Gly
1      5      10      15
Val Gly Tyr Arg Leu Thr His Tyr Glu Lys Lys Ser Leu Lys Leu Phe
20      25      30
Leu Gly Thr Tyr Leu Gly Ser Ser Phe Val Leu Met Leu Val Ile Ser
35      40      45
Val Leu Ala Phe Asn Tyr Glu Lys Asn Glu Lys Ile Lys Xaa Ile Arg

```

173

```

      50              55              60
Met Asp Met Asp Lys Met Ala Ser Lys Ile Ala Ser Glu Ile Ile Gln
65              70              75              80
Leu His Met Gln Thr His Ala Asp Tyr His Asn Ala Leu Asn Ala Leu
      85              90              95
Ile Ser Arg Tyr Lys Asp Val Ser Ile Xaa Leu Xaa Asp Thr
      100             105             110

```

## (2) INFORMATION FOR SEQ ID NO:186:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 124 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...124

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186

```

Leu Met Thr Lys Ser Leu Lys Leu Ile Gln Lys Gly Val Lys Asn Leu
1              5              10              15
Tyr Glu Thr Leu Lys Asn Arg Ala Leu Glu His Gln Asp Thr Leu Met
      20              25              30
Val Gly Arg Ser His Gly Val Phe Gly Glu Pro Ile Thr Phe Gly Leu
      35              40              45
Val Leu Ala Leu Phe Ala Asp Glu Ile Lys Arg His Leu Lys Ala Leu
      50              55              60
Asp Leu Thr Met Glu Phe Ile Xaa Val Gly Ala Ile Ser Gly Ala Met
65              70              75              80
Gly Asn Phe Ala His Ala Pro Leu Glu Leu Glu Glu Leu Ala Cys Gly
      85              90              95
Phe Leu Gly Leu Lys Thr Ala Asn Ile Ser Asn Gln Val Ile Gln Arg
      100             105             110
Asp Arg Tyr Ala Gly Leu His Ala Ile Trp Leu Phe
      115             120

```

## (2) INFORMATION FOR SEQ ID NO:187:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 95 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...95

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187

```

Leu Ser Asp Ala Ser Lys Arg Ser Leu Asn Pro Thr Leu Met Met Asn
1              5              10              15
Asn Asn Asn Thr Leu Pro Lys Pro Leu Glu Glu Ser Leu Asp Leu Lys
      20              25              30

```

SUBSTITUTE SHEET (RULE 26)

174

```

Glu Phe Ile Ala Leu Phe Lys Thr Phe Phe Ala Lys Glu Arg Gly Ser
   35           40           45
Ile Ala Leu Glu Asn Asp Leu Lys Gln Ala Phe Thr Tyr Leu Asn Glu
   50           55           60
Val Asp Ala Ile Gly Leu Pro Ala Pro Xaa Lys Arg Glu Arg Lys Arg
   65           70           75           80
Ser Tyr Cys Cys Gln Thr His Gln Ile Arg Asp Ala Pro Phe Arg
           85           90           95

```

## (2) INFORMATION FOR SEQ ID NO:188:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...80

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:188

```

Leu Pro Ile Ile Leu Xaa Val Ile Val Met Met Phe Phe Ser Lys Ile
1           5           10           15
Val Gly Asp Phe Ile Glu Lys His Tyr Arg Val Lys Thr Leu Ala Phe
           20           25           30
Val Phe Leu Leu Val Val Gly Val Phe Leu Phe Leu Glu Gly Leu His
           35           40           45
Leu His Ile Asn Lys Asn Tyr Leu Tyr Ala Gly Ile Gly Phe Ala Leu
           50           55           60
Leu Ile Glu Cys Leu Xaa Ile Phe Ile Glu Lys Lys Met Lys Lys Ser
           65           70           75           80

```

## (2) INFORMATION FOR SEQ ID NO:189:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 265 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...265

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189

```

Met Ile Lys Ala Arg Phe Lys Lys Arg Leu Leu Gly Ser Arg Gly Ala
1           5           10           15
Phe Asp Leu Asn Ile Asp Leu Glu Ile Lys Glu Ala Glu Val Val Ala
           20           25           30
Leu Leu Gly Glu Ser Gly Ala Gly Lys Ser Thr Ile Leu Arg Ile Leu
           35           40           45
Ala Gly Leu Glu Ala Val Ser Ser Gly Tyr Ile Glu Ala Asn His Ser
           50           55           60
Val Trp Leu Asp Thr Gln Lys Lys Ile Phe Leu Lys Pro Gln Gln Arg

```

175

65		70		75		80									
Lys	Ile	Gly	Phe	Val	Phe	Gln	Asp	Tyr	Ala	Leu	Phe	Pro	His	Leu	Asn
			85						90					95	
Val	Tyr	Gln	Asn	Ile	Ala	Phe	Ala	His	Pro	Lys	Asp	Lys	Asn	Lys	Ile
			100					105					110		
His	Glu	Val	Leu	Arg	Leu	Met	Arg	Leu	Glu	Asn	Leu	Ser	Gln	Gln	Lys
		115					120					125			
Ile	Pro	Lys	Leu	Ser	Gly	Gly	Gln	Ala	Gln	Arg	Val	Ala	Leu	Ala	Arg
		130				135					140				
Ala	Leu	Ile	Ala	Ala	Lys	Asn	Leu	Leu	Leu	Leu	Asp	Glu	Pro	Leu	Asn
		145			150					155				160	
Ala	Leu	Asp	Asn	Ala	Leu	Lys	Asn	Glu	Val	Gln	Gln	Gly	Leu	Leu	Asp
			165						170					175	
Phe	Ile	Lys	Arg	Glu	Asn	Leu	Ser	Val	Leu	Leu	Val	Ser	His	Asp	Pro
			180					185					190		
Asn	Glu	Ile	Thr	Lys	Leu	Ala	Arg	Thr	Phe	Leu	Phe	Leu	Asn	Asn	Gly
		195					200					205			
Val	Ile	Asp	Pro	Asn	Gln	Glu	Asn	Arg	Leu	Phe	Ser	Asn	Arg	Leu	Leu
		210				215					220				
Val	Lys	Pro	Leu	Phe	Glu	Asp	Glu	Asn	Tyr	Cys	His	Tyr	Glu	Val	Ile
		225			230					235				240	
Pro	Gln	Thr	Ile	Ser	Leu	Pro	Lys	Asp	Cys	Leu	Asn	Pro	Thr	Phe	Lys
			245						250					255	
Leu	Asp	Phe	Ile	Gln	Asn	Lys	Lys	Phe							
			260					265							

## (2) INFORMATION FOR SEQ ID NO:190:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 64 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...64

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:190

Val	Lys	Phe	Ser	Val	Leu	Thr	Leu	Phe	Pro	Gln	Leu	Ile	Leu	Pro	Tyr
1				5				10					15		
Phe	Glu	Asp	Ser	Ile	Leu	Lys	Arg	Ala	Leu	Glu	Lys	Asn	Leu	Phe	Glu
		20					25					30			
Leu	Glu	Val	Leu	Asn	Leu	Arg	Asp	Phe	Ser	Ala	Asn	Lys	Tyr	Gln	Lys
		35				40					45				
Ala	Xaa	Ser	His	Ala	His	Trp	Trp	Gly	Cys	Gly	Ala	Asn	Phe	Arg	Pro
	50					55				60					

## (2) INFORMATION FOR SEQ ID NO:191:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 138 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*

176

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...138

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:191

```

Leu Trp Arg Thr Pro Lys Thr Pro Leu Val Ile Lys Pro Tyr Leu Lys
1          5          10          15
Ser Met Ser Asp Ser Glu Ile Phe Ala Xaa Xaa Cys Val Gly Met Ala
          20          25          30
Ser Val Xaa Gly Pro Val Leu Ala Gly Tyr Ala Ser Met Gly Ile Pro
          35          40          45
Leu Pro Tyr Leu Ile Ala Ala Ser Phe Met Ser Ala Pro Gly Gly Leu
50          55          60
Leu Phe Ala Lys Thr Ile Tyr Pro Gln Asn Glu Thr Ile Ser Ser His
65          70          75          80
Ala Asp Val Ser Ala Glu Glu His Val Asn Ile Ile Glu Ala Xaa Ala
          85          90          95
Xaa Gly Ala Ser Thr Gly Xaa His Leu Ala Leu His Val Gly Ala Met
          100          105          110
Leu Leu Ala Phe Val Gly Met Val Ala Leu Val Asn Gly Leu Leu Gly
          115          120          125
Val Val Gly Gly Phe Leu Gly Met Glu His
130          135

```

## (2) INFORMATION FOR SEQ ID NO:192:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 116 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...116

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:192

```

Val Met Asn Phe Phe Val Gly Gly Leu Ser Ile Val Cys Asn Val Val
1          5          10          15
Val Ile Thr Tyr Ser Ala Leu His Pro Thr Ala Pro Val Glu Gly Ala
          20          25          30
Glu Asp Ile Val Gln Val Ser His Leu Thr Ser Phe Tyr Gly Pro
          35          40          45
Ala Thr Gly Leu Leu Phe Gly Phe Thr Tyr Leu Tyr Ala Ala Ile Asn
50          55          60
His Thr Phe Gly Leu Asp Trp Arg Pro Tyr Ser Trp Tyr Ser Leu Phe
65          70          75          80
Val Ala Ile Asn Thr Val Pro Ala Ala Ile Leu Ser His Tyr Ser Asp
          85          90          95
Met Leu Asp Asp His Lys Val Leu Gly Ile Thr Glu Gly Asp Trp Trp
          100          105          110
Ala Ile Ile Xaa
          115

```

## (2) INFORMATION FOR SEQ ID NO:193:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 227 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

177

- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...227
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:193

```

Val Leu Leu Gly Lys His Ser Gly Ala Gly Leu Leu Ser Ala Leu Xaa
1      5      10      15
Ala Leu Ser Phe Gly Ser Gly Val Val Ser Ile Gln Ala Leu Glu Cys
20      25      30
Glu Ile Thr Ser Asn Asn Lys Pro Leu Glu Leu Val Phe Cys Glu Asn
35      40      45
Phe Pro Lys Lys Leu Ser Ala Phe Ala Leu Gly Met Gly Leu Glu Asn
50      55      60
Ile Pro Lys Asp Phe Lys Lys Trp Leu Glu Leu Ala Pro Cys Val Leu
65      70      75      80
Asp Ala Gly Val Phe Tyr His Lys Glu Val Leu Gln Ala Leu Glu Lys
85      90      95
Glu Val Ile Leu Thr Pro His Pro Lys Glu Phe Leu Ser Leu Leu Lys
100     105     110
Ser Val Gly Ile Asn Ile Ser Met Leu Glu Leu Leu Asp Asn Lys Leu
115     120     125
Glu Ile Ala Arg Asp Phe Ser Gln Lys Tyr Pro Lys Val Val Leu Leu
130     135     140
Leu Lys Gly Ala Asn Thr Leu Ile Ala His Gln Gly Arg Val Phe Ile
145     150     155     160
Asn Asn Leu Gly Ser Val Ala Leu Ala Lys Ala Gly Ser Gly Asp Val
165     170     175
Leu Ala Gly Leu Ile Val Ser Leu Leu Ser Gln Asn Tyr Thr Pro Leu
180     185     190
Xaa Ala Ala Ile Asn Ala Ser Leu Ala His Ala Leu Ala Gly Leu Xaa
195     200     205
Phe Lys Asn Xaa Xaa Ala Leu Thr Pro Xaa Asp Leu Ile Glu Lys Xaa
210     215     220
Lys Arg Leu
225

```

(2) INFORMATION FOR SEQ ID NO:194:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 109 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:  
 (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION 1...109
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:194

```

Val Xaa Leu Tyr Leu Ala Leu Thr Leu Ser Leu Gly Ile Ala Met Leu
1      5      10      15
Leu Val Glu Met Leu Ile Gly Asn Leu Gly Lys Lys Asp Val Val Ser
20      25      30
Asn Tyr Gln Ile Leu Asp Pro Lys Arg Lys Lys Tyr Tyr Pro Phe Thr

```

**SUBSTITUTE SHEET (RULE 26)**

178

```

          35          40          45
Ser Phe Phe Ile Leu Gly Gly Pro Leu Ile Leu Ser Phe Tyr Ala Val
   50          55          60
Val Leu Gly Trp Val Leu Tyr Tyr Leu Phe Val Val Thr Phe Asp Leu
   65          70          75          80
Pro Lys Asp Leu Xaa Gln Ala Lys Met Gln Phe Xaa Met Leu Gln Asn
          85          90          95
Gly Ser Leu Ile Trp Pro Val Ile Asp Phe Ser Ala Cys
          100          105

```

## (2) INFORMATION FOR SEQ ID NO:195:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 97 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...97

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:195

```

Leu Thr Thr Lys Ala Cys Trp Leu Leu Arg Val Cys Cys Tyr Arg Ser
1          5          10          15
Leu Asn Ile Thr Ile Lys Asp Arg Thr Met Lys Thr Asn Gly His Phe
          20          25          30
Lys Asp Phe Ala Trp Lys Lys Cys Phe Leu Gly Ala Ser Val Val Ala
          35          40          45
Leu Leu Val Gly Cys Ser Pro His Ile Ile Glu Thr Asn Glu Val Ala
          50          55          60
Leu Lys Leu Asn Tyr His Pro Ala Ser Glu Lys Val Gln Ala Leu Asp
          65          70          75          80
Glu Lys Ile Leu Leu Arg Pro Ala Phe Gln Tyr Ser Xaa Asn Ile
          85          90          95
Cys

```

## (2) INFORMATION FOR SEQ ID NO:196:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 145 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...145

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:196

```

Leu Ser Glu Trp Gln Thr Phe Cys Leu Lys Asp Leu Gly Lys Ile Val
1          5          10          15
Gly Gly Ala Thr Pro Pro Thr Asn Asn Pro Lys Asn Tyr Gly Asn Lys
          20          25          30

```



179

```

Ile Ala Trp Ile Thr Pro Lys Asp Leu Ser Thr Leu Gln Gly Arg Tyr
   35         40         45
Ile Lys Lys Gly Ser Arg Ser Ile Ser Arg Leu Gly Phe Lys Ser Cys
   50         55         60
Ser Cys Val Leu Leu Pro Lys His Ala Ile Leu Phe Ser Ser Arg Ala
  65         70         75         80
Pro Ile Gly Tyr Val Ala Ile Ala Glu Lys Arg Leu Cys Thr Asn Gln
   85         90
Gly Phe Lys Ser Ile Ile Pro Asn Lys Lys Ile Tyr Phe Glu Phe Leu
  100        105        110
Tyr Tyr Leu Leu Lys Tyr Tyr Lys Asp Asn Ile Ser Asn Ile Gly Gly
  115        120        125
Gly Thr Thr Phe Lys Glu Val Ser Gly Ala Thr Leu Gly Ser Ile Pro
  130        135        140
Ser
145

```

## (2) INFORMATION FOR SEQ ID NO:197:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 273 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...273

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197

```

Met Glu Phe Met Lys Lys Phe Val Ala Leu Gly Leu Leu Ser Ala Val
 1         5         10         15
Leu Ser Ser Ser Leu Leu Ala Glu Gly Val Tyr Ile Gly Thr
   20         25         30
Asn Tyr Gln Leu Gly Gln Ala Arg Leu Asn Ser Asn Ile Tyr Asn Thr
  35         40         45
Gly Asp Cys Thr Gly Ser Val Val Gly Cys Pro Pro Gly Leu Thr Ala
  50         55         60
Asn Lys His Asn Pro Gly Gly Thr Asn Ile Asn Trp His Ser Lys Tyr
  65         70         75         80
Ala Asn Gly Ala Leu Asn Gly Phe Gly Leu Asn Val Gly Tyr Lys Lys
   85         90         95
Phe Phe Gln Phe Lys Ser Leu Asp Met Thr Ser Lys Trp Phe Gly Phe
  100        105        110
Arg Val Tyr Gly Leu Phe Asp Tyr Gly His Ala Asp Leu Gly Lys Gln
  115        120        125
Val Tyr Ala Pro Asn Lys Ile Gln Leu Asp Met Val Ser Trp Gly Val
  130        135        140
Gly Ser Asp Leu Leu Ala Asp Ile Ile Asp Lys Asp Asn Ala Ser Phe
  145        150        155        160
Gly Ile Phe Gly Gly Val Ala Ile Gly Gly Asn Thr Trp Lys Ser Ser
  165        170        175
Ala Ala Asn Tyr Trp Lys Glu Gln Ile Ile Glu Ala Lys Gly Pro Asp
  180        185        190
Val Cys Thr Pro Thr Tyr Cys Asn Pro Asn Ala Pro Tyr Ser Thr Asn
  195        200        205
Thr Ser Thr Val Ala Phe Gln Val Trp Leu Asn Phe Gly Val Arg Ala
  210        215        220
Asn Ile Tyr Lys His Asn Gly Val Glu Phe Gly Val Arg Val Pro Leu
  225        230        235        240
Leu Ile Asn Lys Phe Leu Ser Ala Gly Pro Asn Ala Thr Asn Leu Tyr
  245        250        255

```

180

Tyr His Leu Lys Arg Asp Tyr Ser Leu Tyr Leu Gly Tyr Asn Tyr Thr  
                   260                  265                  270  
 Phe

## (2) INFORMATION FOR SEQ ID NO:198:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 148 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...148

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:198

Leu	Val	Gln	Ile	Val	Val	Val	Phe	Tyr	Gly	Leu	Pro	Ala	Leu	Gly	Val
1				5					10					15	
Tyr	Met	Asp	Pro	Ile	Pro	Ala	Gly	Ile	Ile	Ala	Phe	Ser	Phe	Asn	Val
			20					25						30	
Gly	Ala	Tyr	Ala	Ser	Glu	Thr	Leu	Arg	Ala	Ser	Phe	Leu	Ser	Val	Pro
			35					40						45	
Lys	Asp	Gln	Trp	Asp	Ser	Ser	Leu	Ser	Leu	Gly	Leu	Asn	Tyr	Leu	Gln
	50					55					60				
Thr	Phe	Trp	His	Val	Ile	Phe	Phe	Gln	Ala	Leu	Lys	Val	Ala	Thr	Pro
65					70					75				80	
Ser	Leu	Ser	Asn	Thr	Phe	Ile	Ser	Leu	Phe	Lys	Glu	Thr	Ser	Leu	Ala
				85					90					95	
Ser	Val	Val	Thr	Ile	Ala	Glu	Xaa	Phe	Arg	Ile	Ala	Gln	Gln	Lys	Xaa
			100					105					110		
Asn	Val	Ser	Tyr	Asp	Phe	Xaa	Pro	Ile	Tyr	Leu	Glu	Xaa	Ala	Leu	Ile
			115				120					125			
Tyr	Trp	Leu	Phe	Cys	Leu	Val	Leu	Glu	Val	Ile	Gln	Lys	Arg	Val	Glu
	130					135					140				
Lys	Ile	Leu	Asn												
145															

## (2) INFORMATION FOR SEQ ID NO:199:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...134

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:199

Val	Val	Ala	Asp	Glu	Val	Arg	Asn	Leu	Ala	Gly	Arg	Thr	Gln	Lys	Ser
1					5					10				15	
Leu	Ala	Glu	Ile	Asn	Ser	Thr	Ile	Met	Val	Ile	Val	Gln	Glu	Ile	Asn

181

```

                20                25                30
Asp Val Ser Ser Gln Met Asn Leu Asn Ser Gln Lys Met Glu Arg Leu
      35      40      45
Ser Asp Met Ser Lys Ser Val Gln Glu Thr Tyr Glu Lys Met Ser Ser
      50      55      60
Asn Leu Ser Ser Val Val Leu Asp Ser Asn Gln Ser Met Asp Asp Tyr
      65      70      75      80
Ala Lys Ser Gly His Gln Ile Glu Ala Met Val Ser Asp Phe Ala Glu
      85      90      95
Val Glu Lys Val Ala Ser Lys Thr Leu Ala Asp Ser Ser Asp Ile Leu
      100      105      110
Asn Ile Ala Thr His Val Ser Gly Thr Thr Met Asn Leu Xaa Lys Gln
      115      120      125
Val Asn Leu Phe Lys Thr
      130

```

## (2) INFORMATION FOR SEQ ID NO:200:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 133 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...133

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:200

```

Met Asn Tyr Asp Asn Tyr Trp Asp Glu Asp Lys Pro Glu Leu Asn Ile
1      5      10      15
Thr Pro Leu Val Asp Val Met Leu Val Leu Leu Ala Ile Leu Met Val
      20      25      30
Thr Thr Pro Thr Leu Thr Tyr Lys Glu Glu Ile Ala Leu Pro Ser Gly
      35      40      45
Ser Lys Thr Ala Arg Ala Thr Gln Asp Lys Val Ile Glu Ile Arg Met
      50      55      60
Asp Lys Asp Ala Lys Ile Tyr Ile Asp Ser Gln Thr Tyr Glu Tyr Xaa
      65      70      75      80
Ser Phe Pro Asp Thr Phe Asn Leu Leu Ser Lys Lys Tyr Asp Lys Asp
      85      90      95
Thr Arg Val Ser Ile Arg Ala Asp Lys Arg Leu Thr Tyr Asp Lys Val
      100      105      110
Ile Tyr Leu Leu Lys Thr Ile Lys Glu Ala Gly Phe Leu Lys Val Ser
      115      120      125
Leu Ile Thr Ser Pro
      130

```

## (2) INFORMATION FOR SEQ ID NO:201:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 71 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

182

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...71

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:201

```

Met Pro Pro Thr Xaa Pro Gln Ala Ser Ile Leu Arg Leu Thr Leu Lys
1          5          10          15
Asn Pro Leu Xaa Xaa Leu Ser Arg Tyr Ser Leu Cys Leu Leu Lys Lys
          20          25          30
Thr Arg Leu Gln Thr Thr Ser Asn Ser Ala Pro Lys Ala Cys Leu Ile
          35          40          45
Ala Gly Leu Leu Lys Lys Ser Lys Pro Phe Ile Leu Asn Thr Leu Lys
          50          55          60
Ile Arg Ser Leu Leu Lys Pro
65          70

```

## (2) INFORMATION FOR SEQ ID NO:202.

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 217 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...217

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:202

```

Met Pro Val Ile Arg Val Leu Val Met Leu Ala Thr Met Met Met Lys
1          5          10          15
Leu Val Lys Thr Ala Lys Glu Lys Lys Val Phe Lys Asn Val Gly Ile
          20          25          30
Ser Ile Met Gly Ile Ala Phe Trp Glu Ala Ile Lys Asp Ser Ile Lys
          35          40          45
Lys Gln Ile Lys Lys Ser Asp Trp Ile Cys Gly Asn Val Lys Thr Ala
          50          55          60
Asp Asp Tyr Leu Lys Thr His Pro Asn Ser Trp Phe Asn Ser Ala Ile
65          70          75          80
Gly Val Thr Ala Ile Thr Ala Met Leu Met Asn Val Cys Phe Ala Asp
          85          90          95
Asp Gln Ser Lys Lys Glu Val Ala Gln Ala Gln Lys Glu Ala Glu Asn
          100          105          110
Ala Arg Asp Arg Ala Asn Lys Ser Gly Ile Glu Leu Glu Gln Glu Glu
          115          120          125
Gln Lys Thr Glu Gln Glu Lys Gln Lys Thr Glu Gln Glu Lys Gln Lys
          130          135          140
Thr Glu Gln Glu Lys Gln Lys Thr Glu Gln Glu Lys Gln Lys Thr Glu
145          150          155          160
Gln Glu Lys Gln Lys Thr Ser Asn Ile Glu Thr Asn Asn Gln Ile Lys
          165          170          175
Val Glu Gln Glu Gln Gln Lys Thr Glu Gln Glu Lys Gln Lys Thr Asn
          180          185          190
Asn Thr Gln Lys Asp Leu Val Asn Lys Ala Glu Gln Asn Cys Gln Glu
          195          200          205
Asn His Asn Gln Phe Phe Ile Lys Asn
210          215

```

## (2) INFORMATION FOR SEQ ID NO:203:

## (i) SEQUENCE CHARACTERISTICS:

183

(A) LENGTH: 75 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...75

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203

```

Met Val Ile Ser Gly His Phe Thr Thr Tyr Ser Tyr Ile Glu Pro Phe
1      5      10
Ile Ile Gln Ile Ser Gln Phe Ser Pro Asp Ile Thr Thr Leu Met Leu
      20      25      30
Phe Val Phe Gly Leu Ala Gly Val Val Gly Ser Phe Leu Phe Gly Arg
      35      40      45
Leu Tyr Ala Lys Asn Ser Arg Lys Phe Ile Ala Phe Ala Met Val Leu
      50      55      60
Val Ile Cys Pro Gln Pro Leu Ala Phe Cys Val
65      70      75
  
```

(2) INFORMATION FOR SEQ ID NO:204:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 192 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...192

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204

```

Met Lys Ser Thr Arg Ile Gly Ser Lys Ile Val Met Met Val Cys Ala
1      5      10
Val Val Ile Val Ile Ser Ala Val Met Gly Val Ile Ile Ser Tyr Lys
      20      25      30
Val Glu Ser Val Leu Gln Ser Gln Ala Thr Glu Leu Leu Gln Lys Lys
      35      40      45
Ala Gln Leu Val Ser Phe Lys Ile Gln Gly Ile Met Lys Arg Ile Phe
      50      55      60
Met Gly Ala Asn Thr Leu Glu Arg Phe Leu Ser Asp Glu Asn Gly Ala
65      70      75      80
Ile Asn Asp Thr Leu Lys Arg Arg Met Leu Ser Glu Phe Leu Leu Ala
      85      90      95
Asn Pro His Val Leu Leu Val Ser Ala Ile Tyr Thr Asn Asn Asn Glu
      100      105      110
Arg Met Ile Thr Ala Met Asn Met Asp Ser Lys Ile Ala Tyr Pro Asn
      115      120      125
Thr Ala Leu Asn Glu Asn Met Thr Xaa Pro Ile His Ser Leu Lys Ser
      130      135      140
Ile Thr Arg Ser Xaa Pro Tyr Tyr Lys Glu Val Asn Xaa Xaa Lys Ile
      145      150      155      160
Tyr Xaa Xaa Xaa Ile Thr Leu Pro Leu Xaa Xaa Lys Asn Xaa Asn Xaa
  
```

SUBSTITUTE SHEET (RULE 26)

184

		165				170				175
Ile	Xaa	Xaa	Leu	Asn	Phe	Xaa	Leu	Asn	Ile	Asp
		180				185				190
									Xaa	Phe
									Leu	Tyr

## (2) INFORMATION FOR SEQ ID NO:205:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 253 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...253

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:205

Met	Ala	Tyr	Lys	Tyr	Asp	Arg	Asp	Leu	Glu	Phe	Leu	Lys	Gln	Leu	Glu
1				5					10					15	
Ser	Ser	Asp	Leu	Leu	Asp	Leu	Phe	Glu	Val	Leu	Val	Phe	Gly	Lys	Asp
			20					25					30		
Gly	Glu	Lys	Arg	His	Asn	Glu	Lys	Leu	Thr	Ser	Ser	Ile	Glu	Tyr	Lys
		35				40						45			
Arg	His	Gly	Asp	Asp	Tyr	Ala	Lys	Tyr	Ala	Glu	Arg	Ile	Ala	Glu	Glu
	50					55				60					
Leu	Gln	Tyr	Tyr	Gly	Ser	Asn	Ser	Phe	Ala	Ser	Phe	Ile	Lys	Gly	Glu
65					70				75					80	
Gly	Val	Leu	Tyr	Lys	Glu	Ile	Leu	Cys	Asp	Val	Cys	Asp	Lys	Leu	Lys
				85				90					95		
Val	Asn	Tyr	Asn	Lys	Lys	Thr	Glu	Thr	Leu	Ile	Glu	Gln	Asn	Met	
			100				105					110			
Leu	Ser	Lys	Ile	Leu	Glu	Arg	Ser	Leu	Glu	Glu	Met	Asp	Asp	Glu	Glu
		115					120					125			
Val	Lys	Glu	Met	Cys	Asp	Glu	Leu	Ser	Ile	Lys	Asn	Thr	Asp	Asn	Leu
	130					135					140				
Asn	Arg	Gln	Ala	Leu	Ser	Ala	Ala	Thr	Leu	Thr	Leu	Phe	Lys	Met	Gly
145					150					155					160
Gly	Phe	Lys	Ser	Tyr	Gln	Leu	Ala	Val	Ile	Val	Ala	Asn	Ala	Val	Ala
				165					170					175	
Lys	Thr	Ile	Leu	Gly	Arg	Gly	Leu	Ser	Leu	Ala	Gly	Asn	Gln	Val	Leu
			180				185						190		
Thr	Arg	Thr	Leu	Ser	Phe	Leu	Thr	Gly	Pro	Val	Gly	Trp	Ile	Ile	Thr
		195					200					205			
Gly	Val	Trp	Thr	Ala	Ile	Asp	Ile	Ala	Gly	Pro	Ala	Tyr	Arg	Val	Thr
	210					215					220				
Ile	Pro	Ala	Cys	Ile	Val	Val	Ala	Thr	Leu	Arg	Leu	Lys	Thr	Gln	Gln
225					230					235					240
Ala	Asn	Glu	Asp	Lys	Lys	Ser	Leu	Gln	Ile	Glu	Ser	Val			
				245					250						

## (2) INFORMATION FOR SEQ ID NO:206:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 293 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:

185

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...293

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:206

```

Leu Leu Leu Phe Ile Val Val Ile Thr Ser Leu Val Lys Asn Thr Ile
1      5      10      15
Pro Asn Ile Trp Leu Thr Lys Ile Leu Tyr Met Ala Ile Leu Leu Cys
20      25      30
Ala Ile Ala His Ser Val Gly Xaa Ile Leu Arg Trp Tyr Val Ser Gly
35      40      45
His Ser Pro Trp Ser Asn Ala Tyr Glu Ser Met Phe Tyr Ile Ala Trp
50      55      60
Ala Ser Val Ile Ala Gly Phe Val Leu Arg Xaa Lys Leu Ala Leu Ser
65      70      75      80
Ala Ser Ser Phe Leu Ala Gly Ile Ala Leu Phe Val Ala His Leu Gly
85      90      95
Phe Met Asp Pro Gln Ile Gly Pro Leu Val Pro Val Leu Lys Ser Tyr
100     105     110
Trp Leu Asn Ile His Val Ser Val Ile Thr Ala Ser Tyr Gly Phe Leu
115     120     125
Gly Leu Cys Phe Val Leu Gly Ile Leu Ser Leu Val Leu Phe Ile Leu
130     135     140
Arg Lys Gln Gly Arg Phe Asn Leu Asp Lys Thr Ile Leu Ser Ile Ser
145     150     155     160
Ala Ile Asn Glu Met Ser Met Ile Leu Gly Leu Phe Met Leu Thr Ala
165     170     175
Gly Asn Phe Leu Gly Gly Val Trp Ala Asn Glu Ser Trp Gly Arg Tyr
180     185     190
Trp Gly Trp Asp Pro Lys Glu Thr Trp Ala Leu Ile Ser Ile Cys Val
195     200     205
Tyr Ala Leu Ile Leu His Leu Arg Phe Leu Gly Ser Gln Asn Trp Pro
210     215     220
Phe Ile Leu Ala Ser Ser Ser Val Leu Gly Phe Tyr Ser Val Leu Met
225     230     235     240
Thr Leu Phe Trp Arg Glu Leu Leu Pro Phe Trp Leu Ala Gln Leu Cys
245     250     255
Arg Arg Xaa Ser Phe Ala Asp Pro Tyr Phe Phe Ile Leu Phe Gly Ser
260     265     270
Asp Thr Phe Arg Ser Arg Ile Leu Ala Tyr Phe Lys Arg His Leu Ser
275     280     285
Leu Pro Lys Leu Val
290

```

## (2) INFORMATION FOR SEQ ID NO:207:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 142 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...142

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:207

```

Val Glu Met Ile His Thr Gln Asp Tyr Ile Lys Met Glu Glu Ala Ala
1      5      10      15

```

186

```

Thr Glu Ala Ile Lys Arg Lys Glu Ser Ser Ile Tyr Leu Gly Met Asp
      20                      25          30
Ile Leu Lys Asn Gly Ala Asp Ala Leu Ile Ser Ala Gly His Ser Gly
      35                      40          45
Ala Thr Met Gly Leu Ala Thr Leu Arg Leu Gly Arg Ile Lys Gly Val
      50                      55          60
Glu Arg Pro Ala Ile Cys Thr Leu Met Pro Ser Val Gly Lys Arg Pro
      65                      70          75          80
Ser Val Leu Leu Asp Ala Gly Ala Asn Thr Asp Cys Lys Pro Glu Tyr
      85                      90          95
Leu Ile Asp Phe Ala Leu Met Gly Tyr Glu Tyr Ala Lys Ser Val Leu
      100                     105          110
His Tyr Asp Ser Pro Lys Val Gly Leu Leu Ser Asn Gly Glu Glu Asp
      115                     120          125
Ile Lys Gly Gly Ile Arg Ser Leu Lys Lys Arg Ile Lys Cys
      130                     135          140

```

## (2) INFORMATION FOR SEQ ID NO:208:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...144

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:208

```

Met Leu Glu Ile Lys Asn Leu Asn Cys Val Leu Asn Ser His Phe Ser
1      5      10      15
Leu Gln Asn Ile Asn Ile Ser Leu Ser Tyr Ser Glu Arg Val Ala Ile
      20      25      30
Val Gly Glu Ser Gly Ser Gly Lys Ser Ser Ile Ala Asn Leu Val Met
      35      40      45
Arg Leu Asn Pro Arg Phe Lys Ser His Asn Gly Glu Ile Leu Phe Glu
      50      55      60
Thr Thr Asn Leu Leu Lys Glu Ser Glu Ala Phe Xaa Gln His Leu Arg
      65      70      75      80
Gly Asn Ile Ile Ala Tyr Ile Ala Gln Asp Pro Leu Ser Ser Leu Asn
      85      90      95
Pro Leu His Lys Ile Gly Lys Gln Met Ser Glu Ala Tyr Phe Leu His
      100     105     110
His Lys Asn Ala Ser Gln Val Ser Leu Asn Glu Gln Val Leu Asn Val
      115     120     125
Met Lys Gln Val Gln Leu Asp Glu Asn Phe Trp Asn Val Ser Leu Met
      130     135     140

```

## (2) INFORMATION FOR SEQ ID NO:209:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 83 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*



187

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...83

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:209

```

Met Asn Tyr Lys Val Ala Ser Ala Arg Asn Ile Ala Thr Leu Leu Phe
1          5          10          15
Leu Phe Phe Ser Gln Ser Glu Ala Phe Asp Leu Gly Lys Ile Ala Lys
          20          25          30
Ile Lys Ala Gly Ala Glu Ser Phe Ser Lys Val Gly Phe Asn Asn Lys
          35          40          45
Pro Ile Asn Xaa Asn Lys Gly Ile Tyr Pro Thr Glu Thr Phe Met Thr
          50          55          60
Ile Asn Gly Leu His Ala Gly Gly Phe Tyr Gly Ala Leu Ala Gln Lys
65          70          75          80
Arg Tyr Gly

```

## (2) INFORMATION FOR SEQ ID NO:210:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...130

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:210

```

Met Asp Ala Leu Glu Ile Thr Gln Lys Leu Ile Ser Tyr Pro Thr Ile
1          5          10          15
Thr Pro Lys Glu Cys Gly Ile Phe Glu Tyr Ile Lys Ser Leu Phe Pro
          20          25          30
Ala Phe Lys Thr Leu Glu Cys Glu Lys Asn Gly Val Lys Asn Leu Phe
          35          40          45
Leu Tyr Arg Ile Phe Asn Pro Leu Lys Lys His Ala Glu Lys Glu His
          50          55          60
Ala Lys Glu Lys His Val Lys Glu Asn Val Xaa Pro Leu His Phe Cys
65          70          75          80
Xaa Ala Gly His Ile Xaa Val Val Pro Pro Gly Xaa Xaa Xaa Xaa
          85          90          95
Asp Ser Phe Xaa Xaa Ile Ile Lys Glu Gly Phe Leu Tyr Gly Arg Gly
          100          105          110
Ala Gln Asp Met Lys Gly Gly Val Gly Xaa Phe Xaa Arg Cys Xaa Xaa
115          120          125
Lys Phe
130

```

## (2) INFORMATION FOR SEQ ID NO:211:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 340 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

SUBSTITUTE SHEET (RULE 26)

188

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...340

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:211

```

Met Ile Leu Ser Ile Glu Ser Ser Cys Asp Asp Ser Ser Leu Ala Leu
1      5      10      15
Thr Arg Ile Glu Asp Ala Lys Leu Ile Ala His Phe Lys Ile Ser Gln
      20      25      30
Glu Lys His His Ser Ser Tyr Gly Gly Val Val Pro Glu Ile Ala Ser
      35      40      45
Arg Leu His Ala Glu Asn Leu Pro Leu Leu Leu Glu Arg Val Lys Ile
      50      55      60
Ser Leu Asn Lys Asp Phe Ser Lys Ile Lys Ala Ile Ala Ile Thr Asn
      65      70      75      80
Gln Pro Gly Leu Ser Val Thr Leu Ile Glu Gly Leu Met Met Ala Lys
      85      90      95
Ala Leu Ser Leu Ser Leu Asn Leu Pro Leu Ile Leu Glu Asp His Leu
      100      105      110
Arg Gly His Val Tyr Ser Leu Phe Ile Asn Glu Lys Gln Thr Arg Met
      115      120      125
Pro Leu Ser Val Leu Leu Val Ser Gly Gly His Ser Leu Ile Leu Glu
      130      135      140
Ala Arg Asp Tyr Glu Asp Ile Lys Ile Val Ala Thr Ser Leu Asp Asp
      145      150      155      160
Ser Phe Gly Glu Ser Phe Asp Lys Val Ser Lys Met Leu Asp Leu Gly
      165      170      175
Tyr Pro Gly Gly Pro Ile Val Glu Lys Leu Ala Leu Asp Tyr Ala His
      180      185      190
Pro Asn Glu Pro Leu Met Phe Pro Ile Pro Leu Lys Asn Ser Pro Asn
      195      200      205
Leu Ala Phe Ser Phe Ser Gly Leu Lys Asn Ala Val Arg Leu Glu Val
      210      215      220
Glu Lys Asn Ala His Asn Leu Asn Asp Glu Val Lys Gln Lys Ile Gly
      225      230      235      240
Tyr His Phe Gln Ser Ala Ala Ile Glu His Leu Ile Gln Gln Thr Lys
      245      250      255
Arg Tyr Phe Lys Ile Lys Arg Pro Lys Ile Phe Gly Ile Val Gly Gly
      260      265      270
Ala Ser Gln Asn Leu Ala Leu Arg Lys Ala Phe Glu Asp Leu Cys Ala
      275      280      285
Glu Phe Asp Cys Glu Leu Val Leu Ala Pro Leu Glu Phe Cys Ser Asp
      290      295      300
Asn Ala Ala Met Ile Gly Arg Ser Ser Leu Glu Ala Tyr Gln Lys Lys
      305      310      315      320
Arg Phe Ile Pro Leu Glu Lys Ala Asp Ile Ser Pro Arg Thr Leu Leu
      325      330      335
Lys Asn Phe Glu
      340

```

## (2) INFORMATION FOR SEQ ID NO:212:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 168 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

189

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...168

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:212

```

Met Leu Ser Ser Asn Asp Leu Phe Met Val Val Leu Gly Ala Ile Leu
1      5      10      15
Leu Val Leu Val Cys Leu Val Gly Tyr Leu Tyr Leu Lys Glu Lys Glu
20      25      30
Phe Tyr His Lys Met Arg Arg Leu Glu Lys Thr Leu Asp Glu Ser Tyr
35      40      45
Gln Glu Asn Tyr Leu Tyr Ser Lys Arg Leu Arg Glu Leu Glu Gly Arg
50      55      60
Leu Glu Gly Leu Ser Leu Glu Lys Ser Ala Lys Glu Asp Ser Ser Leu
65      70      75      80
Lys Thr Thr Leu Ser His Leu Tyr Asn Gln Leu Gln Glu Ile Gln Lys
85      90      95
Ser Met Asp Lys Glu Arg Asp Tyr Leu Glu Glu Lys Ile Ile Xaa Xaa
100      105      110
Lys Thr Xaa Xaa Lys Thr Trp Gly Ile Met Pro Leu Ala Met Lys Ser
115      120      125
Thr Glu Lys Gln Val Leu Lys Met Tyr Gln Glu Gly Tyr Ser Val Asp
130      135      140
Ser Ile Ser Lys Glu Phe Lys Val Ser Lys Gly Glu Val Glu Phe Ile
145      150      155      160
Leu Asn Met Ala Gly Leu Lys Trp
165

```

## (2) INFORMATION FOR SEQ ID NO:213:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 121 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...121

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:213

```

Leu Asp Pro Phe Ser His Lys Glu Asn Phe Leu Ala Val Glu Thr Phe
1      5      10      15
Lys Met Leu Gly Lys Thr Glu Ser Lys Asp Asn Leu Asn Trp Met Ile
20      25      30
Ala Leu Ile Ile Glu Lys Asp Lys Val Tyr Glu Gln Val Gly Ser Val
35      40      45
Arg Phe Val Val Val Val Ala Ser Ala Ile Met Val Leu Ala Leu Ile
50      55      60
Ile Ala Ile Thr Leu Leu Met Arg Ala Ile Val Ser Asn Arg Leu Glu
65      70      75      80
Val Val Ser Ser Thr Leu Ser His Phe Phe Lys Leu Leu Asn Asn Gln
85      90      95
Xaa His Ser Ser Xaa Xaa Lys Leu Val Xaa Ala Arg Ser Asn Asp Glu
100      105      110
Leu Gly Arg Xaa Gln Thr Xaa Asp Xaa
115      120

```

## (2) INFORMATION FOR SEQ ID NO:214:

## (i) SEQUENCE CHARACTERISTICS:

SUBSTITUTE SHEET (RULE 26)

190

(A) LENGTH: 149 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...149

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214

```

Met Glu Phe Tyr Gln Val Tyr Asp Pro Leu Gly His Ile Trp Leu Ser
1      5      10      15
Ala Leu Val Ala Leu Ser Pro Ile Ala Leu Phe Phe Ile Ser Leu Ile
20     25     30
Val Phe Lys Leu Lys Gly Tyr Ser Ala Gly Phe Leu Ser Leu Ala Leu
35     40     45
Ser Ile Leu Ile Ala Leu Phe Val Tyr Lys Met Pro Val Gln Met Val
50     55     60
Ser Ala Ser Phe Phe Tyr Gly Phe Leu Tyr Gly Leu Trp Pro Ile Ala
65     70     75     80
Trp Ile Val Ile Ala Ile Phe Leu Tyr Asn Leu Ser Val Lys Ser
85     90     95
Gly Tyr Phe Glu Ile Leu Lys Glu Ser Ile Leu Ser Leu Thr Pro Asp
100    105    110
His Arg Ile Leu Val Ile Leu Ile Gly Phe Cys Phe Gly Ser Phe Leu
115    120    125
Xaa Gly Ala Xaa Gly Phe Gly Gly Pro Val Ala Ile Thr Ala Ala Ile
130    135    140
Leu Val Ala Leu Gly
145
  
```

(2) INFORMATION FOR SEQ ID NO:215:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 325 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...325

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:215

```

Met Lys Arg Ile Leu Val Ser Leu Ala Val Leu Ser His Ser Ala His
1      5      10      15
Ala Val Lys Thr His Asn Leu Glu Arg Val Glu Ala Ser Gly Val Ala
20     25     30
Asn Asp Lys Glu Ala Pro Leu Ser Trp Arg Ser Lys Glu Val Arg Asn
35     40     45
Tyr Met Gly Ser Arg Thr Val Ile Ser Asn Lys Gln Leu Thr Lys Ser
50     55     60
Ala Asn Gln Ser Ile Glu Glu Ala Leu Gln Asn Val Pro Gly Val His
65     70     75     80
Ile Arg Asn Ser Thr Gly Ile Gly Ala Val Pro Ser Ile Ser Ile Arg
  
```

191

[illegible]

(2) INFORMATION FOR SEQ ID NO:216:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 252 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...252

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:216

Leu	Arg	Ser	Ile	Ser	Arg	Ile	Lys	Met	Leu	Ser	Val	Tyr	Glu	Lys	Gly
1				5					10					15	
Asn	Ala	Leu	Asp	Lys	Arg	Val	Leu	Glu	Glu	Trp	Leu	Leu	Ser	Glu	Asp
			20					25					30		
Ile	Leu	Met	Glu	Asn	Ala	Ala	Met	Ala	Leu	Glu	Arg	Ala	Val	Leu	Gln
		35					40					45			
Asn	Ala	Ser	Leu	Gly	Ala	Lys	Val	Ile	Ile	Leu	Cys	Gly	Ser	Gly	Asp
	50					55					60				
Asn	Gly	Gly	Asp	Gly	Tyr	Thr	Leu	Ala	Arg	Arg	Leu	Val	Gly	Arg	Phe
65					70					75				80	
Lys	Thr	Leu	Val	Phe	Glu	Met	Lys	Leu	Ala	Lys	Ser	Pro	Met	Cys	Gln
				85					90					95	
Leu	Gln	Lys	Glu	Arg	Ala	Lys	Lys	Val	Gly	Val	Val	Ile	Lys	Ala	Trp
			100					105					110		
Glu	Glu	Lys	Asn	Glu	Asp	Leu	Glu	Cys	Asp	Val	Leu	Val	Asp	Cys	Val
		115					120					125			
Val	Gly	Ser	Ala	Phe	Lys	Gly	Gly	Leu	Glu	Pro	Phe	Leu	Asp	Phe	Glu

192

130	135	140
Ser Leu Ser Gln Lys Ala Arg Phe Lys Ile Ala Cys Asp Ile Pro Ser		
145	150	155
Gly Ile Asp Ser Lys Gly Arg Val Asp Lys Arg Ala Phe Lys Xaa Gly		160
	165	170
Tyr Arg Leu Ser Ala Trp Ala Leu Phe Lys Ser Cys Leu Leu Ser Xaa		175
	180	185
Lys Xaa Lys Xaa Tyr Ile Xaa Xaa Leu Lys Xaa Xaa His Leu Xaa Val		190
	195	200
Phe Asn Gln Ile Tyr Glu Ile Pro Thr Xaa Thr Phe Leu Leu Glu Lys		205
	210	215
Xaa Asp Leu Lys Leu Pro Leu Arg Asp Arg Lys Lys Arg Ser Gln Arg		220
225	230	235
Arg Leu Arg Ala Cys Ala Cys Ala Phe Gly Gln Ala		240
	245	250

## (2) INFORMATION FOR SEQ ID NO:217:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...138

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:217

Met	Ala	Leu	Asp	Lys	Arg	Ile	Trp	Met	His	Phe	Asp	Leu	Leu	Pro	Phe
1				5					10					15	
Val	Phe	Ile	Ile	Pro	Leu	Leu	Val	Val	Ser	Phe	Leu	Leu	Ile	Phe	Glu
			20					25					30		
Ser	Ser	Ala	Val	Leu	Ser	Leu	Lys	Gln	Gly	Val	Tyr	Tyr	Ala	Ile	Gly
		35					40				45				
Phe	Leu	Leu	Phe	Trp	Val	Val	Phe	Phe	Ile	Pro	Phe	Arg	Lys	Leu	Asp
	50					55				60					
Arg	Trp	Leu	Phe	Ala	Leu	Tyr	Trp	Ala	Cys	Val	Ile	Leu	Leu	Ala	Leu
65				70						75				80	
Val	Asp	Phe	Met	Gly	Ser	Ser	Lys	Leu	Gly	Ala	Gln	Arg	Trp	Leu	Val
			85					90						95	
Ile	Pro	Phe	Thr	Ser	Ile	Thr	Leu	Gln	Pro	Ser	Glu	Pro	Val	Lys	Asn
		100					105						110		
Arg	Xaa	Ser	Phe	Ile	Val	Gly	Ala	Phe	Xaa	Xaa	Asn	Xaa	Pro	Asp	Xaa
		115				120						125			
Leu	Leu	Arg	Ala	Met	Ile	Gly	Ala	Cys	Phe						
	130					135									

## (2) INFORMATION FOR SEQ ID NO:218:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 326 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

193

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...326

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:218

```

Val Leu Met Ala Leu Xaa Asp Lys Arg Tyr Gly Leu Glu Ala Gly Ile
1      5      10      15
Lys Tyr Phe Thr Met Gly Ala Met Ala Ser Ala Phe Phe Ala Met Gly
      20      25      30
Ala Met Ala Phe Tyr Leu Leu Thr Gly Ser Leu Asn Leu Glu Val Ile
      35      40      45
Thr Leu Tyr Leu His Thr Glu Gly Ile Thr Asn Pro Met Leu Phe Ala
      50      55      60
Met Gly Thr Ile Phe Leu Ile Gly Ala Ile Gly Phe Lys Val Ser Leu
      65      70      75      80
Val Pro Phe His Thr Trp Met Pro Asp Val Tyr Glu Gly Asn Asn Pro
      85      90      95
Val Phe Ala Ser Tyr Ile Ser Ile Val Pro Lys Ile Ala Gly Phe Val
      100      105      110
Val Ala Thr Arg Leu Phe Gly Ala Phe Ile Asp Thr His Thr Ala Trp
      115      120      125
Val Glu Asp Ile Phe Tyr Val Leu Ile Leu Met Thr Ile Thr Ile Pro
      130      135      140
Asn Phe Ile Ala Leu Trp Gln Glu Asp Val Lys Arg Met Leu Ala Tyr
      145      150      155      160
Ser Ser Ile Ser His Ser Gly Phe Ala Leu Ala Cys Val Phe Ile His
      165      170      175
Thr Glu Asp Ser Gln Gln Ala Met Phe Val Tyr Trp Phe Met Phe Ala
      180      185      190
Phe Thr Tyr Ile Gly Ala Phe Gly Leu Leu Trp Leu Leu Lys Ser Arg
      195      200      205
Glu Lys Thr Trp Asp Glu Arg Tyr Asp His Pro Tyr Ser Lys Phe Asn
      210      215      220
Gly Leu Ile Lys Thr His Pro Leu Val Ala Ile Leu Gly Ala Ile Phe
      225      230      235      240
Val Phe Gly Leu Ala Gly Ile Pro Pro Phe Ser Val Phe Trp Gly Lys
      245      250      255
Phe Leu Ala Val Glu Ser Ala Leu Glu Ser Asn His Ile Leu Leu Ala
      260      265      270
Val Val Met Leu Val Asn Ser Ala Val Ala Ala Phe Tyr Phe Arg
      275      280      285
Trp Leu Val Ala Met Phe Phe Asn Lys Pro Leu Gln Thr Gln Ser Tyr
      290      295      300
Ala Lys Thr Ile Phe Thr Pro Lys Thr Pro Pro Cys Pro Phe Met Arg
      305      310      315      320
Ser Leu Leu Pro Trp Arg
      325

```

## (2) INFORMATION FOR SEQ ID NO:219:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 240 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...240

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:219

194

```

Met Ile Asn Ser Lys Lys Ser Leu Lys Lys Gly Leu Arg Gly Phe Phe
1      5      10      15
Lys Ile Leu Lys Asp Arg Asn Gly Ala His Phe Ser Cys Gly Ala Thr
20      25      30
Ser Gly Phe Gly Leu Glu Ile Ala Lys Ala Phe Leu Gln Lys Asn His
35      40      45
Val Val Phe Gly Thr Gly Arg Arg Gln Glu Asn Leu Gln Lys Leu Gln
50      55      60
Leu Ala Tyr Pro Lys Arg Phe Ile Pro Leu Cys Phe Asp Leu Gln Asn
65      70      75      80
Lys Pro Glu Thr Lys Arg Ala Ile Glu Thr Ile Phe Ser Met Thr Asp
85      90      95
Arg Ile Asp Ala Leu Ile Asn Asn Ala Gly Leu Ala Leu Gly Leu Asn
100     105     110
Lys Ala Tyr Glu Cys Glu Leu Asp Asp Trp Glu Val Met Ile Asp Thr
115     120     125
Asn Ile Lys Gly Leu Leu His Leu Thr Arg Leu Ile Leu Pro Ser Met
130     135     140
Ile Glu His Asp Gln Gly Thr Ile Ile Asn Leu Gly Ser Ile Ala Gly
145     150     155     160
Thr Tyr Ala Tyr Pro Gly Gly Xaa Val Tyr Gly Ala Ser Lys Ala Xaa
165     170     175
Val Lys Gln Xaa Ser Xaa Asn Leu Arg Ala Asp Val Ala Gly Thr Asn
180     185     190
Thr Arg Gly Arg Arg Trp Asn Pro Gly Cys Val Ala Lys Pro Lys Val
195     200     205
Ser Arg Val Arg Gly Lys Gly Asp Lys Pro Lys Pro Lys Ser Gly Tyr
210     215     220
Glu Lys His Pro Leu Pro Gln Thr Thr Arg Gln Gly Leu Thr Ser Gly
225     230     235     240

```

## (2) INFORMATION FOR SEQ ID NO:220:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 204 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...204
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:220

```

Val Ser Gly Val Val Leu Ser Lys Phe Asp Ser Asp Ser Lys Gly Gly
1      5      10      15
Ile Ala Leu Gly Ile Thr Tyr Gln Leu Gly Leu Pro Leu Arg Phe Ile
20      25      30
Gly Ser Gly Glu Lys Ile Pro Asp Leu Asp Val Phe Met Pro Glu Arg
35      40      45
Ile Val Gly Arg Leu Met Gly Ala Gly Asp Ile Ile Ser Leu Ala Glu
50      55      60
Lys Thr Ala Ser Val Leu Asn Pro Asn Glu Ala Lys Asp Leu Ser Lys
65      70      75      80
Lys Leu Lys Lys Gly Gln Phe Thr Phe Asn Asp Phe Leu Asn Gln Ile
85      90      95
Glu Lys Val Lys Lys Leu Gly Ser Met Ser Ser Leu Ile Ser Met Ile
100     105     110
Pro Gly Leu Gly Asn Met Ala Ser Ala Leu Lys Asp Thr Asp Leu Glu
115     120     125
Ser Ser Leu Glu Val Lys Lys Ile Lys Ala Met Val Asn Ser Met Thr
130     135     140

```



195

```

Lys Lys Glu Arg Glu Asn Pro Glu Ile Leu Asn Gly Ser Arg Arg Lys
145      150      155      160
Arg Ile Ala Leu Gly Xaa Gly Leu Glu Xaa Xaa Glu Ile Asn Arg Ile
      165      170      175
Ile Lys Arg Phe Asp Gln Ala Ser Lys Met Ala Lys Arg Leu Thr Asn
      180      185      190
Lys Lys Gly Ile Ser Asp Leu Met Asn Leu Xaa Xaa
      195      200

```

## (2) INFORMATION FOR SEQ ID NO:221:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 92 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...92

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:221

```

Val Glu Lys Ala His Pro Asp Val Phe Asn Leu Leu Leu Gln Val Leu
1      5      10      15
Asp Glu Gly His Leu Thr Asp Ser Lys Gly Val Arg Val Asp Phe Lys
      20      25      30
Asn Thr Ile Leu Ile Leu Thr Ser Asn Val Ala Ser Gly Ala Leu Leu
      35      40      45
Glu Glu Asp Leu Ser Glu Ala Asp Lys Gln Lys Ala Ile Lys Glu Ser
      50      55      60
Leu Arg Gln Phe Phe Lys Pro Glu Phe Leu Asn Arg Leu Asp Glu Ile
      65      70      75      80
Ile Ser Phe Asn Ala Leu Asp Ser His Ala Ile Ile
      85      90

```

## (2) INFORMATION FOR SEQ ID NO:222:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 82 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...82

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:222

```

Leu Val Phe Leu Asp Arg Arg Leu Ile Val Met Val Thr Asp Ser Lys
1      5      10      15
Gly Ser Arg Tyr Ile Asn Val His Ile Leu Phe Arg Gln Ile Ser Leu
      20      25      30
Tyr Ala Leu Leu Ser Val Val Gly Ser Leu Leu Phe Leu Gly Val Ser
      35      40      45
Leu Leu Val Leu Asn Lys Glu Ile Lys Asn Ile Glu Lys Gln His Ala

```

196

```

      50              55              60
Leu Xaa Thr Lys Glu Phe Glu Lys Lys Arg Glu Thr Asn Glu Xaa Leu
65              70              75              80
Ser Xaa

```

## (2) INFORMATION FOR SEQ ID NO:223:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 233 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...233

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:223

```

Leu Ser Leu Met Xaa Val Leu Asn Ala Lys Glu Cys Val Xaa Pro Ile
1      5      10      15
Thr Arg Ser Val Lys Tyr His Gln Gln Ser Ala Glu Ile Arg Ala Leu
20      25      30
Gln Leu Gln Ser Tyr Lys Met Ala Lys Met Ala Leu Asp Asn Asn Leu
35      40      45
Lys Leu Val Lys Asp Lys Lys Pro Ala Val Ile Leu Asp Leu Asp Glu
50      55      60
Thr Val Leu Asn Thr Phe Asp Tyr Ala Gly Tyr Leu Val Lys Asn Cys
65      70      75      80
Ile Lys Tyr Thr Pro Glu Thr Trp Asp Lys Phe Glu Lys Glu Gly Ser
85      90      95
Leu Thr Leu Ile Pro Gly Ala Leu Asp Phe Leu Glu Tyr Ala Asn Ser
100      105      110
Lys Gly Val Lys Ile Phe Tyr Ile Ser Asn Arg Thr Gln Lys Asn Lys
115      120      125
Ala Phe Thr Leu Lys Thr Leu Lys Ser Phe Lys Leu Pro Gln Val Ser
130      135      140
Glu Glu Ser Val Leu Leu Lys Glu Lys Gly Lys Pro Lys Ala Val Arg
145      150      155      160
Arg Glu Leu Val Ala Lys Asp Tyr Ala Ile Val Leu Gln Val Gly Asp
165      170      175
Thr Leu His Asp Phe Asp Ala Ile Phe Ala Lys Asp Ala Lys Asn Ser
180      185      190
Gln Glu Gln Gln Ala Lys Val Leu Gln Asn Ala Gln Lys Phe Gly Thr
195      200      205
Glu Trp Ile Ile Leu Pro Asn Ser Leu Tyr Gly Thr Trp Glu Asp Gly
210      215      220
Pro Ile Lys Ala Trp Gln Asn Lys Lys
225      230

```

## (2) INFORMATION FOR SEQ ID NO:224:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 85 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

197

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...85

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:224

```

Met Leu Ala Ala Gly Leu Thr Leu Pro Glu Phe Gly Cys Tyr Leu Ser
1      5      10      15
His Tyr Leu Leu Trp Lys Glu Cys Val Lys Leu Asp Gln Pro Val Val
      20      25      30
Ile Leu Glu Asp Asp Val Thr Leu Glu Ser His Phe Met Gln Ala Leu
      35      40      45
Glu Asp Cys Leu Lys Ser Pro Phe Asp Phe Val Arg Leu Tyr Gly Cys
      50      55      60
Tyr Trp Tyr Tyr Gln Arg Asp Lys Ile Pro Cys Phe Ala Gln Arg Ile
65      70      75      80
Cys Ile Ser Ser Leu
      85

```

## (2) INFORMATION FOR SEQ ID NO:225:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 115 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...115

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:225

```

Leu Ile Ala Leu Arg Val Thr Ala Trp Lys Val Xaa Ala Met Lys Arg
1      5      10      15
Leu His Leu Ser Val Lys Asp Ala Glu Asn Phe Asp Ala Ile Leu Arg
      20      25      30
Glu Arg Pro Phe Phe Lys Asp Leu Ile Glu Phe Met Val Ser Gly Pro
      35      40      45
Val Val Val Met Val Leu Glu Gly Lys Asp Ala Val Ala Lys Asn Arg
      50      55      60
Glu Leu Met Gly Ala Thr Asp Pro Lys Leu Ala Gln Lys Gly Thr Ile
65      70      75      80
Arg Ala Asp Phe Ala Glu Ser Ile Asp Ala Asn Ala Val His Gly Ser
      85      90      95
Asp Ser Leu Glu Asn Ala His Asn Glu Ile Ala Phe Phe Phe Ala Ala
100      105      110
Arg Glu Phe
      115

```

## (2) INFORMATION FOR SEQ ID NO:226:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 394 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

SUBSTITUTE SHEET (RULE 26)

198

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...394

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:226

```

Leu Met Trp Leu Lys Thr Leu Thr Leu Gln Thr Leu Asn Thr Asp Lys
1      5      10      15
Ala Leu Gln Glu Phe Ser Lys Thr Met Glu Ala Phe Lys Thr Lys Leu
20      25      30
Ile Gln Ser Ala Asn Asp Val His Ser Glu Thr Ser Arg Ala Ala Ile
35      40      45
Ala Asn Asp Leu Glu Arg Leu Lys Glu His Met Ile Asn Val Ala Asn
50      55      60
Thr Ser Ile Gly Gly Glu Phe Leu Phe Gly Gly Ser Lys Val Asp Arg
65      70      75      80
Pro Pro Ile Asp Ser Asn Gly Lys Tyr His Gly Asn Gly Glu Asp Leu
85      90      95
Asn Ala Leu Ile Ser Ser Asp Asn Leu Val Pro Tyr Asn Ile Ser Gly
100     105     110
Gln Asp Leu Phe Leu Gly Thr Asp Lys Asp Lys His Lys Leu Ile Thr
115     120     125
Thr Asn Ile Lys Leu Leu Asn Gln Asn Lys Leu Xaa Pro Asp Val Met
130     135     140
Asp Ala Leu Glu His Ser Ser Leu Pro Glu Glu Val Phe Ile Lys Pro
145     150     155     160
Ser Asp Thr Leu Arg Glu Leu Ile Gly Asp Asn Asp Lys Asn Pro Thr
165     170     175
Asn Asp Pro Lys Glu Phe Phe Tyr Leu Gln Gly Ile Arg Pro Asp Gly
180     185     190
Ser Ser Phe Lys Glu Lys Phe Ala Leu Asp Lys Ala Tyr Gln Asn Gln
195     200     205
Glu Ser Ala Thr Lys Val Ser Asp Leu Leu Asp Lys Ile Gly His Ala
210     215     220
Tyr Gly Asn Thr Ser Gln Asn Lys Val Val Asp Val Ser Leu Asn Asn
225     230     235     240
Trp Gly Gln Ile Glu Ile Lys Asn Leu Thr Pro Gly Ser Glu Asn Leu
245     250     255
Asp Phe His Leu Ile Ser Ser Asp Gly Asp Phe Asp Asp Leu Asp Ala
260     265     270
Leu Arg Ser Ser Gly Lys Arg Val Thr Glu Tyr Val Lys Ser Ala Phe
275     280     285
Val Thr Asp Arg Ser Leu Ser Gln Val Lys Ala Val Pro Asn Met Tyr
290     295     300
Asn Pro Lys Val Leu Glu Ile Pro Ser Val Phe Val Thr Lys Asp Asn
305     310     315     320
Val Leu Ala Asn Lys Asn Thr Lys Leu Ser Glu Ile Phe Gly Asp Lys
325     330     335
Val Glu Thr Leu Lys Ile Asn Ala Ser Arg Leu Gly Asp Glu Ser Ala
340     345     350
Ile Lys Ile Pro Asn Leu Pro Ile Asn Leu Asp Ile Pro Ile Leu Leu
355     360     365
Asp Val Lys Asn Ser Thr Ile Lys Asp Leu Lys Asp Ala Ile Lys Glu
370     375     380
Arg Phe Asn Asn Glu Gly Gly Cys Gly Asn
385     390

```

## (2) INFORMATION FOR SEQ ID NO:227:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

199

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...102

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227

```

Leu Lys Ala Leu Asn Asp Cys Met Val Phe Phe His Lys Lys Ile Ile
1      5      10
Leu Asn Phe Ile Tyr Ser Leu Met Val Ala Phe Leu Phe His Leu Ser
20     25     30
Tyr Gly Val Leu Leu Lys Ala Asp Gly Met Ala Lys Lys Gln Thr Leu
35     40     45
Leu Val Gly Glu Arg Leu Val Trp Asp Lys Leu Thr Leu Leu Gly Phe
50     55     60
Leu Glu Lys Asn His Ile Pro Gln Lys Leu Tyr Tyr Asn Leu Ser Ser
65     70     75     80
Gln Asp Lys Glu Leu Ser Ala Glu Ile Gln Ser Asn Val Thr Tyr Tyr
85     90     95
Xaa Phe Lys Arg Cys Lys
100

```

(2) INFORMATION FOR SEQ ID NO:228:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 363 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...363

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:228

```

Met Lys Phe Phe Leu Leu Lys Lys Phe Ser Xaa Phe Leu Asn Thr Gln
1      5      10
Thr His Phe Asn Leu Lys Arg Leu Asn Ala Ser Ser Phe Leu Leu Glu
20     25     30
Thr Phe Ser Lys Glu Lys His Ala Phe Val Val Asp Leu Ser Ala Pro
35     40     45
Tyr Ile Gly Leu Ser Lys Lys Pro Pro Glu Ser Val Leu Lys Asn Thr
50     55     60
Leu Ala Leu Asp Phe Cys Leu Asn Lys Phe Thr Lys Asn Ala Lys Ile
65     70     75     80
Leu Gln Ala Asn Val Ile Asp Asn Asp Arg Ile Leu Glu Ile Lys Gly
85     90     95
Ala Lys Asp Leu Ala Tyr Lys Ser Glu Thr Phe Ile Leu Arg Leu Glu
100    105    110
Met Ile Pro Lys Lys Ala Asn Leu Met Ile Leu Asp Gln Glu Lys Cys
115    120    125
Val Ile Glu Ala Phe Arg Phe Asn Asp Arg Val Ala Lys Asn Asp Ile
130    135    140
Leu Gly Ala Leu Pro Pro Asn Ile Tyr Glu His Gln Glu Glu Asp Leu
145    150    155    160
Asp Phe Lys Gly Leu Leu Asp Ile Leu Glu Lys Asp Phe Leu Ser Tyr
165    170    175
Gln His Lys Glu Leu Glu His Lys Lys Asn Gln Ile Ile Lys Arg Leu

```

SUBSTITUTE SHEET (RULE 26)

200

			180					185					190			
Asn	Ala	Gln	Lys	Glu	Arg	Leu	Lys	Glu	Lys	Leu	Glu	Lys	Leu	Glu	Asp	
		195					200					205				
Pro	Lys	Thr	Leu	Gln	Leu	Glu	Ala	Lys	Glu	Leu	Gln	Thr	Gln	Ala	Ser	
	210					215					220					
Leu	Leu	Leu	Thr	Tyr	Gln	His	Leu	Ile	Asn	Arg	Arg	Glu	Asn	Arg	Val	
225					230					235					240	
Ile	Leu	Lys	Asp	Phe	Glu	Asp	Lys	Glu	Cys	Met	Ile	Glu	Ile	Asp	Lys	
			245						250					255		
Ser	Met	Pro	Leu	Asn	Ala	Phe	Ile	Asn	Lys	Lys	Phe	Thr	Leu	Ser	Lys	
			260					265					270			
Lys	Lys	Lys	Gln	Lys	Ser	Gln	Phe	Leu	Tyr	Leu	Glu	Glu	Glu	Asn	Leu	
		275					280					285				
Lys	Glu	Lys	Ile	Ala	Phe	Lys	Glu	Asn	Gln	Ile	Asn	Tyr	Val	Arg	Asp	
		290				295					300					
Ala	Ala	Glu	Glu	Ser	Val	Leu	Glu	Met	Phe	Met	Pro	Val	Lys	Asn	Ser	
305					310				315						320	
Lys	Ile	Lys	Arg	Pro	Met	Asn	Gly	Tyr	Glu	Val	Leu	Tyr	Tyr	Lys	Asp	
				325				330						335		
Xaa	Lys	Xaa	Gly	Leu	Gly	Lys	Thr	Lys	Lys	Arg	Ile	Ser	Ser	Phe	Tyr	
			340					345					350			
Lys	Thr	Gln	Xaa	Arg	Met	Ile	Xaa	Gly	Cys	Xaa						
		355					360									

(2) INFORMATION FOR SEQ ID NO:229:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- ```

FEATURE
(A) NAME/KEY: misc_feature
(B) LOCATION 1...22

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:229

ATATCCATGG TGAGTTTGAT GA

22

(2) INFORMATION FOR SEQ ID NO:230:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...25

201

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:230  
ATGAATTCAA TTTTTTATTT TGCCA 25

(2) INFORMATION FOR SEQ ID NO:231:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:231  
AATTCCATGG TGGGGGCTAT G 21

(2) INFORMATION FOR SEQ ID NO:232:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:232  
ATGAATTCTC GATAGCCAAA ATC 23

(2) INFORMATION FOR SEQ ID NO:233:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

202

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:233

ATTTCATGG TCATGTCTCA TATT

24

(2) INFORMATION FOR SEQ ID NO:234:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:234

ATGAATTCCA TCTTTTATTC CAC

23

(2) INFORMATION FOR SEQ ID NO:235:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:  
(A) ORGANISM: Helicobacter pylori

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:235

AACCATGGTG ATTTTAAGCA TTGAAAG

27

(2) INFORMATION FOR SEQ ID NO:236:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 28 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular



203

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...28
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:236

AAGAATTCCA CTCAAATTT TTTAACAG

28

## (2) INFORMATION FOR SEQ ID NO:237:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...25
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:237

GATCATCCAT ATGTTATCTT CTAAT

25

## (2) INFORMATION FOR SEQ ID NO:238:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...23
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:238

TGAATTCAAC CATTTTAACC CTG

23

## (2) INFORMATION FOR SEQ ID NO:239:

204

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...27
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:239

TATACCATGG TGAAATTTTT TCTTTTA

27

## (2) INFORMATION FOR SEQ ID NO:240:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...25
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:240

AGAATTCAAT TGCGTCTTGT AAAAG

25

## (2) INFORMATION FOR SEQ ID NO:241:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...25

205

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:241

TTATGGATCC AAACCAATTA AAAC

25

(2) INFORMATION FOR SEQ ID NO:242:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:242

TATCTCGAGT TATAGAGAAG GGC

23

(2) INFORMATION FOR SEQ ID NO:243:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...22

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:243

ATATCCATGG TGAGTTTGAT GA

22

(2) INFORMATION FOR SEQ ID NO:244:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

206

(A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:244

ATGAATTCAA TTTTATTT TGCCA

25

## (2) INFORMATION FOR SEQ ID NO:245:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245

AATTCATGG CTATCCAAAT CCG

23

## (2) INFORMATION FOR SEQ ID NO:246:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Helicobacter pylori

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246

ATGAATTCGC CAAAATCGTA GTATT

25

## (2) INFORMATION FOR SEQ ID NO:247:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

207

- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...24
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:247

GATACCATGG AATTTATGAA AAAG

24

## (2) INFORMATION FOR SEQ ID NO:248:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...25
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:248

TGAATTCGAA AAAGTGTAGT TATAC

25

## (2) INFORMATION FOR SEQ ID NO:249:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...22
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:249

TTGAACACTT TTGATTATGC GG

22

## (2) INFORMATION FOR SEQ ID NO:250:

- (i) SEQUENCE CHARACTERISTICS:

208

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...23
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:250

GGATTATGCG ATTGTTTTAC AAG

23

## (2) INFORMATION FOR SEQ ID NO:251:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...21
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:251

GTCTTTAGCA AAAATGGCGT C

21

## (2) INFORMATION FOR SEQ ID NO:252:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...21
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:252

AATGAGCGTA AGAGAGCCTT C

21

## (2) INFORMATION FOR SEQ ID NO:253:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:253

CTTATGGGGG TATTGTCA

18

## (2) INFORMATION FOR SEQ ID NO:254:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:254

AGCATGTGGG TATCCAGC

18

## (2) INFORMATION FOR SEQ ID NO:255:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori

210

- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:255  
AGGTTGTTGC CTAAAGACT

19

(2) INFORMATION FOR SEQ ID NO:256:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori

- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256  
CTGCCTCCAC CTTTGATC

18

(2) INFORMATION FOR SEQ ID NO:257:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori

- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:257  
ACCAATATCA ATTGGCACT

19

(2) INFORMATION FOR SEQ ID NO:258:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)



211

- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...18
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:258

ACTTGGAAGA GCTCTGCA

18

## (2) INFORMATION FOR SEQ ID NO:259:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...19
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:259

CTTGCTTGTC ATATCTAGC

19

## (2) INFORMATION FOR SEQ ID NO:260:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: *Helicobacter pylori*
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...18
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:260

GTTGAAGTGT TGGTGCTA

18

## (2) INFORMATION FOR SEQ ID NO:261:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs

212

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...22
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:261

CAAGCAAGTG GTTTGGTTTT AG

22

## (2) INFORMATION FOR SEQ ID NO:262:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...22
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:262

TGGAAAGAGC AAATCATTGA AG

22

## (2) INFORMATION FOR SEQ ID NO:263:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...21
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:263

213

GCCCATATATC AAAAAGCCCA T

21

## (2) INFORMATION FOR SEQ ID NO:264:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:264

CTAAAACCAA ACCACTTGCT TGTC

24

## (2) INFORMATION FOR SEQ ID NO:265:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 16 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION 1...16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:265

GTAAAACGAC GGCCAG

16

## (2) INFORMATION FOR SEQ ID NO:266:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Helicobacter pylori*

214

- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...17

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:266

CAGGAAACAG CTATGAC

17

(2) INFORMATION FOR SEQ ID NO:267:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:267

ATCTTACCTA TCACCTCAAA T

21

(2) INFORMATION FOR SEQ ID NO:268:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 21 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: circular
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Helicobacter pylori
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION 1...21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:268

AGACAGCAAC ATCTTTGTGA A

21

-215-

5    CLAIMS

1. A substantially pure nucleic acid encoding an *H. pylori* polypeptide, said nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NOs:1-114.

10

2. A substantially pure nucleic acid from naturally occurring *H. pylori* which hybridizes under stringent conditions to a nucleic acid which encodes an *H. pylori* polypeptide, said nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:1-114.

15

3. A method of evaluating a compound for the ability to bind an *H. pylori* nucleic acid comprising: contacting said compound with an *H. pylori* nucleic acid selected from the group consisting of SEQ ID NO:1-114 and determining if said compound binds said *H. pylori* nucleic acid.

20

4. The method of claim 3, wherein said compound is an activator of the bacterial life cycle.

5. The method of claim 3, wherein said compound is an inhibitor of the bacterial life cycle.

25

6. The method of claim 3, wherein said method is performed *in vitro*.

7. The method of claim 3, wherein said method is performed *in vivo*.

30

8. A method of generating a vaccine for immunizing a subject against *H. pylori* infection comprising: immunizing said subject with a nucleic acid encoding an *H. pylori* polypeptide or a fragment thereof, said nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:1-114, and a therapeutically acceptable carrier.

35

9. A method of detecting the presence of a *Helicobacter* species in a sample comprising:

contacting said sample with a nucleic acid encoding an *H. pylori* polypeptide, said nucleic acid comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:1-114;

40

hybridizing said sample to said nucleic acid;

-216-

5 said hybridization being indicative of the presence of said *Helicobacter* species in said sample.

10. The method of claim 9, wherein said *Helicobacter* species is *H. pylori*.

10 11. The method of claim 9, wherein said nucleic acid is 20 or more nucleotides in length.

12. A method of inhibiting expression of a gene from a *Helicobacter* species comprising: administering to said species an *H. pylori* antisense nucleic acid selected from  
15 the group consisting of SEQ ID NO:1-114.

13. The method of claim 12, wherein said *Helicobacter* species is *H. pylori*.

14. The method of claim 12, wherein said antisense nucleic acid is administered in  
20 a carrier.

15. The method of claim 12, wherein said carrier is a liposome or a bacteriophage.

16. The method of claim 12, wherein said antisense nucleic acid is 20 or more  
25 nucleotides in length.

17. The method of claim 12, wherein said antisense nucleic acid is capable of binding to *Helicobacter* nucleic acid or mRNA.

18. A method of generating a vaccine for immunizing a subject against *H. pylori* infection comprising: immunizing said subject with a nucleic acid encoding an *H. pylori* polypeptide or a fragment thereof, said polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:115-228, and a therapeutically acceptable carrier.

35

| Sequence Identifier | Sequence Name | BLAST hit                      | Description                                               |
|---------------------|---------------|--------------------------------|-----------------------------------------------------------|
| 1.                  | 3987580.      |                                |                                                           |
| 2.                  | 55843.        |                                |                                                           |
| 3.                  | 1365943.      |                                |                                                           |
| 4.                  | 914087.       |                                |                                                           |
| 5.                  | 2343887.      |                                |                                                           |
| 6.                  | 24409641.     |                                |                                                           |
| 7.                  | 26258562.     |                                |                                                           |
| 8.                  | 5138.         |                                |                                                           |
| 9.                  | 21647676.     |                                |                                                           |
| 10.                 | 207817.       |                                |                                                           |
| 11.                 | 116018.       |                                |                                                           |
| 12.                 | 486075.       |                                |                                                           |
| 13.                 | 30708287.     |                                |                                                           |
| 14.                 | 6828218.      |                                |                                                           |
| 15.                 | 24089087.     |                                |                                                           |
| 16.                 | 35163962.     |                                |                                                           |
| 17.                 | 6288949.      |                                |                                                           |
| 18.                 | 35345228.     |                                |                                                           |
| 19.                 | 24406567.     |                                |                                                           |
| 20.                 | 24409577.     |                                |                                                           |
| 21.                 | 15126875.     |                                |                                                           |
| 22.                 | 25595387.     |                                |                                                           |
| 23.                 | 5869090.      |                                |                                                           |
| 24.                 | 23912807.     |                                |                                                           |
| 25.                 | 598933.       |                                |                                                           |
| 26.                 | 24500088.     |                                |                                                           |
| 27.                 | 4882842.      |                                |                                                           |
| 28.                 | 4062813.      |                                |                                                           |
| 29.                 | 35269000.     |                                |                                                           |
| 30.                 | 23535937.     |                                |                                                           |
| 31.                 | 2042312.      |                                |                                                           |
| 32.                 | 30478562.     |                                |                                                           |
| 33.                 | 34161500.     |                                |                                                           |
| 34.                 | 33203192.     |                                |                                                           |
| 35.                 | 12505125.     |                                |                                                           |
| 36.                 | 22379952.     |                                |                                                           |
| 37.                 | 489057.       |                                |                                                           |
| 38.                 | 5312712.      |                                |                                                           |
| 39.                 | 24220627.     |                                |                                                           |
| 40.                 | 12698442.     |                                |                                                           |
| 41.                 | 4338438.      |                                |                                                           |
| 42.                 | 2149041.      |                                |                                                           |
| 43.                 | 4569693.      |                                |                                                           |
| 44.                 | 3179505.      |                                |                                                           |
| 45.                 | 33397538.     |                                |                                                           |
| 46.                 | 917152.       |                                |                                                           |
| 47.                 | 34172639.     |                                |                                                           |
| 48.                 | 30730068.     |                                |                                                           |
| 49.                 | 23631292.     |                                |                                                           |
| 50.                 | 3962777.      |                                |                                                           |
|                     |               | sp P20021 CADA_STAAU.(H7P 296) | probable cadmium-transporting ATPase,                     |
|                     |               | . (H7P 149)                    | mature-parasite-infected erythrocyte surface antigen,     |
|                     |               | (H7P 525)                      |                                                           |
|                     |               | (H7P 420)                      | flagellar biosynthetic protein,                           |
|                     |               | sp P10408 SECA_ECOLI.(H7P 243) | norepinephrine transporter,                               |
|                     |               |                                | Protein secretion secA subunit,                           |
|                     |               | sp P26276 ALCC_PSEAE.(H7P 214) | phosphomannomutase,                                       |
|                     |               | sp Q03523 MURE_BACSU.(H7P 194) | UDP-N-ACETYLMURAMYL-TRIPETIDE SYNTHETASE,                 |
|                     |               | sp P31548 YABJ_ECOLI.(H7P 372) | HYPOTHETICAL ABC TRANSPORTER,                             |
|                     |               | sp P18783 EXBB_ECOLI.(H7P 161) | biopolymer transport exbB protein,                        |
|                     |               | sp P31547 YAEF_ECOLI.(H7P 445) | HYPOTHETICAL 23.3 KD PROTEIN-INTEGRAL MEMBRANE,           |
|                     |               | g1 311022 sp L08012 (H7P 232)  | major surface LPS-antigen,                                |
|                     |               | sp P31219 YBBA_ECOLI.(H7P 455) | hypothetical abc transporter n tesa region,               |
|                     |               | (H7P 403)                      | Cell division inhibitor,                                  |
|                     |               | sp P33916 YEFJ_ECOLI.(H7P 348) | HYPOTHETICAL ABC TRANSPORTER,                             |
|                     |               | sp W82917 WOLF_LAG_1.(H7P 553) | minor flagellin flag precursor-H.pylori,                  |
|                     |               | sp P33024 .                    | similar to E.coli hypothetical nucleoside transport prot, |
|                     |               | sp P31122 YDEA_ECOLI.          | similar to CHL/PAMPHEMICOL RESISTANCE PROTEIN,            |

FIGURE 1

|      |           |                                |                                                                   |   |
|------|-----------|--------------------------------|-------------------------------------------------------------------|---|
| 51.  | 24215,    | sp P19933 GLTS_ECOLI. (HP 227) | sodium/glutamate symport carrier protein.                         | 2 |
| 52.  | 3964593,  |                                |                                                                   |   |
| 53.  | 3991067,  |                                |                                                                   |   |
| 54.  | 24410643, |                                |                                                                   |   |
| 55.  | 47290919, |                                |                                                                   |   |
| 56.  | 10742963, |                                |                                                                   |   |
| 57.  | 16422591, |                                |                                                                   |   |
| 58.  | 23490686, | sp P28573 NTPR_RAT. (HP 544)   | SODIUM-DEPENDENT PROLINE TRANSPORTER,                             |   |
| 59.  | 875042,   | gp U09005 VPU09005_4 (HP 144)  | channel component of the sodium-type flagellar motor.             |   |
| 60.  | 4562712,  |                                |                                                                   |   |
| 61.  | 23594818, |                                |                                                                   |   |
| 62.  | 2150290,  |                                |                                                                   |   |
| 63.  | 30471091, | sp P08083 . (HP 436)           | encodes the serologically diverse protein H in Streptococci       |   |
| 64.  | 4821082,  |                                |                                                                   |   |
| 65.  | 23621317, | sp P03819 KEFC_ECOLI. (HP 113) | potassium efflux system protein.                                  |   |
| 66.  | 19531291, | sp P16676 CYSA_ECOLI. (HP 424) | SULFATE TRANSPORT ATP-BINDING,                                    |   |
| 67.  | 36573502, | sp P26093 HEL_HAEIN. (HP 540)  | outer membrane protein P4 precursor.                              |   |
| 68.  | 7116626,  |                                |                                                                   |   |
| 69.  | 12617677, |                                |                                                                   |   |
| 70.  | 34495938, | sp P33916 YEF_ECOLI. (HP 123)  | HYPOTHETICAL ABC TRANSPORTER.                                     |   |
| 71.  | 24218968, | . (HP 226)                     | methyl-accepting chemotaxis protein; transmembrane receptor       |   |
| 72.  | 24634750, | . (HP 162)                     | probable cadmium-transporting ATPase,                             |   |
| 73.  | 24122293, | gp Z31376 BSFLDST_3. (HP 433)  | flagellar protein flis.                                           |   |
| 74.  | 20173437, |                                |                                                                   |   |
| 75.  | 22411050, |                                |                                                                   |   |
| 76.  | 3942217,  |                                |                                                                   |   |
| 77.  | 12520952, |                                |                                                                   |   |
| 78.  | 31681556, |                                |                                                                   |   |
| 79.  | 3907042,  |                                |                                                                   |   |
| 80.  | 24222885, |                                |                                                                   |   |
| 81.  | 17497107, |                                |                                                                   |   |
| 82.  | 19556290, |                                |                                                                   |   |
| 83.  | 34427317, | sp P30750 ABC_ECOLI. (HP 19)   | PREPROTEIN TRANSLOCASE; SECA SUBUNIT,<br>ATP-BINDING PROTEIN ABC. |   |
| 84.  | 11132778, |                                |                                                                   |   |
| 85.  | 24645837, |                                |                                                                   |   |
| 86.  | 35887,    |                                |                                                                   |   |
| 87.  | 21487501, | sp P13511 CZCA_ALCEU. (HP 141) | cation efflux system membrane protein czca,                       |   |
| 88.  | 33601578, | sp P33913 YEA_ECOLI. (HP 301)  | homology to HYPOTHETICAL PROTEIN IN BCR 5' REGION (FRAGMENT)      |   |
| 89.  | 26355190, | sp P37732 MODD_AZOVI. (HP 371) | molybdenum transport atp-binding protein.                         |   |
| 90.  | 3319687,  | gp U05670 HU05670_2. (HP 542)  | influenzae type B lipooligosaccharide,                            |   |
| 91.  | 7225666,  | sp P13511 CZCA_ALCEU. (HP 431) | cation efflux system membrane protein czca,                       |   |
| 92.  | 4826401,  |                                |                                                                   |   |
| 93.  | 867183,   | gi 495471 gp U07145 . (HP 147) | vacuolating cytotoxin of Hpylori.                                 |   |
| 94.  | 21573938, |                                |                                                                   |   |
| 95.  | 21563752, |                                |                                                                   |   |
| 96.  | 2111040,  |                                |                                                                   |   |
| 97.  | 10037799, |                                |                                                                   |   |
| 98.  | 23437502, |                                |                                                                   |   |
| 99.  | 32462543, |                                |                                                                   |   |
| 100. | 20976500, |                                |                                                                   |   |
| 101. | 103812,   | gp Z12001 BFCRP_1. (HP 64)     | chloramphenicol resistance protein.                               |   |
| 102. | 14496077, |                                |                                                                   |   |
| 103. | 4714375,  |                                |                                                                   |   |

FIGURE 1 CONT'D



FIGURE 1 CONT'D

4

|      |           |                                 |                                                                   |
|------|-----------|---------------------------------|-------------------------------------------------------------------|
| 157. | 3261306.  | sp P28635 VAEC_ECOLI, (HPP 91)  | outer membrane 30K protein,                                       |
| 158. | 16225006. | sp P30858 ARTP_ECOLI, (HPP 108) | ARGININE TRANSPORT ATP-BINDING PROTEIN                            |
| 159. | 33595708. |                                 |                                                                   |
| 160. | 17787558. |                                 |                                                                   |
| 161. | 24396937. |                                 |                                                                   |
| 162. | 35360843. | sp U05676 HFU05676.2, (HPP 342) | vacuolating cytotoxin,                                            |
| 163. | 32609403. | sp P10740 , (HPP 366)           | PHOSPHATIDYLSELINE DECARBOXYLASE,                                 |
| 164. | 32705252. |                                 |                                                                   |
| 165. | 429192.   | sp P33979 , (HPP 170)           | FLAGELLAR P-RING PROTEIN PRECURSOR,                               |
| 166. | 22692187. | sp Q03203 , (HPP 11)            | INTEGRAL MEMBRANE PROTEIN,                                        |
| 167. | 10096666. | sp P37169 MVIN_SALTY, (HPP 15)  | VIRULENCE FACTOR MVIN,                                            |
| 168. | 19536458. |                                 |                                                                   |
| 169. | 5194840.  |                                 |                                                                   |
| 170. | 3906963.  | sp P14900 HURD_ECOLI, (HPP 136) | UDP-N-ACETYLMURAMOYLALANINE --D-GLUTAMATELIGASE,                  |
| 171. | 23486342. |                                 |                                                                   |
| 172. | 19536375. |                                 |                                                                   |
| 173. | 17089217. | sp P13511 CZCA_ALCEU, (HPP 204) | cation efflux system proteins,                                    |
| 174. | 23635968. | sp Q02114 CMLB_BACSU, (HPP 294) | N-ACETYLMURAMOYL-L-ALANINE AMIDASE CELL WALL HYDROLASE AUTOLYSIN, |
| 175. | 25992137. |                                 |                                                                   |
| 176. | 14855461. |                                 |                                                                   |
| 177. | 45914063. | sp P35528 FLIP_BACSU, (HPP 104) | flagellar biosynthetic protein flip,                              |
| 178. | 16984442. |                                 |                                                                   |
| 179. | 12120938. |                                 |                                                                   |
| 180. | 23439055. |                                 |                                                                   |
| 181. | 32636635. |                                 |                                                                   |
| 182. | 3933437.  |                                 |                                                                   |
| 183. | 24328910. | sp P13036 FECA_ECOLI, (HPP 362) | IRON(III) DICITRATE TRANSPORT PROTEIN FECA PRECURSOR,             |
| 184. | 29479681. |                                 |                                                                   |
| 185. | 4177212.  |                                 |                                                                   |
| 186. | 26351567. |                                 |                                                                   |
| 187. | 20415937. | sp P21458 SEFE_BACSU, (HPP 219) | spoIIIE gene product,                                             |
| 188. | 24003758. | (HPP 269)                       | iron(III) transport system,                                       |
| 189. | 23853185. |                                 |                                                                   |
| 190. | 26052137. |                                 |                                                                   |
| 191. | 23468781. |                                 |                                                                   |
| 192. | 14726542. |                                 |                                                                   |
| 193. | 6523427.  |                                 |                                                                   |
| 194. | 13704718. |                                 |                                                                   |
| 195. | 23440814. |                                 |                                                                   |
| 196. | 14642217. | sp P37734 MODB_AZOVI, (HPP 515) | molybdate-binding periplasmic protein precursor,                  |
| 197. | 5875152.  | sp P15932 FLCK_SALTY, (HPP 5)   | FLAGELLAR HOOK-ASSOCIATED PROTEIN 1 HAP1,                         |
| 198. | 104792.   | sp Q03523 MURE_BACSU, (HPP 343) | UDP-N-ACETYLMURAMYL-TRIPTEPTIDE SYNTHETASE,                       |
| 199. | 34265691. | sp P07176 PAL_ECOLI, (HPP 572)  | peptidoglycan-associated lipoprotein,                             |
| 200. | 5440436.  | sp P23452 , (HPP 222)           | component of flagellum,                                           |
| 201. | 24078837. |                                 |                                                                   |
| 202. | 12343763. | sp P15509411.                   | spoIIIE gene product,                                             |
| 203. | 19626250. |                                 |                                                                   |
| 204. | 4728193.  |                                 |                                                                   |
| 205. | 1416312.  |                                 |                                                                   |
| 206. | 31413433. |                                 |                                                                   |
| 207. | 26758437. |                                 |                                                                   |
| 208. | 10553192. |                                 |                                                                   |
| 209. | 14480927. |                                 |                                                                   |

FIGURE 1 CONT'D

5

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 12897656,<br>260941,<br>2242337,<br>21486677,<br>6933202,<br>11924177,<br>3166040,<br>3360130,<br>2915903,<br>203192,<br>36203402,<br>26261040,<br>23492181,<br>14570443,<br>22453166,<br>34573431,<br>10407625,<br>35442513,<br>24256572,<br>26301059,<br>11723593,<br>23945317,<br>25995917,<br>26197187,<br>22164962,<br>32627125,<br>16412593,<br>32453958,<br>16459375,<br>3906712,<br>32595137,<br>16440842,<br>31250333,<br>4708337,<br>11253,<br>26306340,<br>1367157,<br>35704718,<br>423131,<br>186752,<br>24230058,<br>24238762,<br>24276587,<br>29557266,<br>43490713,<br>16251627,<br>23915877,<br>4960952,<br>25925,<br>23880087,<br>6093906,<br>29302003,<br>13726562, | sp P31438 YICK_ECOLI,<br>(HPP 140)<br><br>sp Q01960 FLHF_BACSU,<br>(HPP 128)<br>sp P02913 ,<br>(HPP 470)<br>sp U13166 RMU13166_3,<br>(HPP 247)<br>sp P33650 FE0B_ECOLI,<br>(HPP 185)<br><br>gp M31827 BACDDSA_2,<br>(HPP 295)<br><br>gp X70039 HPCAI_1,<br>(HPP 100)<br>sp Q05605 EXBB_PSEPU,<br>(HPP 427)<br><br>g1 531265 gp D21131 ,<br>(HPP 342)<br>sp P17952 ,<br>(HPP 20)<br><br>sp P19933 GLTS_ECOLI,<br>(HPP 110)<br>sp P17448 KCTP_ECOLI,<br>(HPP 233)<br>(HPP 457)<br><br>sp P13036 FECA_ECOLI,<br>(HPP 494)<br><br>sp P33941 YCVI_ECOLI,<br>(HPP 54) | penicillin-binding protein 2,<br><br>flagellar biosynthesis protein flhF,<br>integral membrane protein,<br>chemotaxis protein cheY,<br>iron(II) transport system,<br><br>cell division and sporulation protein,<br><br>cytotoxicity associated immunodominant antigen (H. pylori),<br>biopolymer transport exbD protein,<br><br>sequence predicts membrane bound protein,<br>N-ACETYLPHOSPHATE--ALANINE LIGASE,<br><br>sodium/glutamate symport carrier protein,<br>alpha-ketoglutarate permease,<br><br>glycerolphosphate auxotrophy in plsB background,<br><br>iron dicitrate transport protein,<br><br>HYPOTHETICAL ABC TRANSPORTER. |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

FIGURE 1 CONT'D

6/15

|      |           |                       |           |                                                                         |   |
|------|-----------|-----------------------|-----------|-------------------------------------------------------------------------|---|
| 263, | 5891412,  | sp P23282 ,           | (HPP 518) | 3-deoxy-D-manno-octulosonil: acid transferase,                          | 6 |
| 264, | 2986577,  |                       |           |                                                                         |   |
| 265, | 3906937,  | g1 495471 gp U07145 , | (HPP 26)  | vacuolating cytotoxin of <i>H. pylori</i> ,                             |   |
| 266, | 2458267,  | sp P13738 NHAA_ECOLI, | (HPP 275) | NA+/H+ ANTIporter [E.coli],                                             |   |
| 267, | 24824087, |                       |           |                                                                         |   |
| 268, | 31218912, | sp L36317 YSCCC2A_1,  | (HPP 40)  | Cu+++-transporting P-type ATPase,                                       |   |
| 269, | 3466680,  |                       |           |                                                                         |   |
| 270, | 783432,   |                       |           |                                                                         |   |
| 271, | 24609431, |                       |           |                                                                         |   |
| 272, | 6495137,  |                       |           |                                                                         |   |
| 273, | 4035783,  |                       |           |                                                                         |   |
| 274, | 35417942, |                       |           |                                                                         |   |
| 275, | 24414687, |                       |           |                                                                         |   |
| 276, | 259665,   |                       |           |                                                                         |   |
| 277, | 23867207, |                       |           |                                                                         |   |
| 278, | 34109763, |                       |           |                                                                         |   |
| 279, | 487750,   |                       |           |                                                                         |   |
| 280, | 24411011, | sp P36175 GCP_PASHA,  | (HPP 47)  | O-SIALYLGLYCOPROTEIN ENDOPEPTIDASE-lacks signal sequence,               |   |
| 281, | 4721061,  | sp P10324 PAL_HAEIN,  | (HPP 51)  | OUTER MEMBRANE PROTEIN P6 PRECURSOR                                     |   |
| 282, | 5440436,  | sp P14788 ,           | (HPP 234) | SULFATE TRANSPORT ATP-BINDING PROTEIN CYSA,                             |   |
| 283, | 24329712, |                       |           |                                                                         |   |
| 284, | 10181942, |                       |           |                                                                         |   |
| 285, | 24036302, |                       |           |                                                                         |   |
| 286, | 677088,   |                       |           |                                                                         |   |
| 287, | 5111308,  |                       |           |                                                                         |   |
| 288, | 29456178, |                       |           |                                                                         |   |
| 289, | 10677187, |                       |           |                                                                         |   |
| 290, | 422937,   |                       |           |                                                                         |   |
| 291, | 21511555, | sp P10121 FTSV_ECOLI, | (HPP 251) | CELL DIVISION PROTEIN -- FUNCTIONAL HOMOLOG OF SRP RECEPTOR,            |   |
| 292, | 1370202,  |                       |           |                                                                         |   |
| 293, | 36335436, | sp P10089 HLY2_ECOLI, | (HPP 410) | HARMOlysin SECRETION ATP-BINDING PROTEIN,                               |   |
| 294, | 1962590,  | sp P15928 FLIF_SALTY, | (HPP 114) | FLAGELLAR M-RING PROTEIN,                                               |   |
| 295, | 21976637, |                       |           |                                                                         |   |
| 296, | 24427340, |                       |           |                                                                         |   |
| 297, | 114505,   |                       |           |                                                                         |   |
| 298, | 35428912, |                       |           |                                                                         |   |
| 299, | 19557055, |                       |           |                                                                         |   |
| 300, | 6517040,  |                       |           |                                                                         |   |
| 301, | 35837767, | sp P19933 GLTS_ECOLI, | (HPP 314) | sodium/glutamate symport carrier protein,                               |   |
| 302, | 289711,   |                       |           |                                                                         |   |
| 303, | 21720017, |                       |           |                                                                         |   |
| 304, | 1256885,  |                       |           |                                                                         |   |
| 305, | 25422192, |                       |           |                                                                         |   |
| 306, | 719606,   |                       |           |                                                                         |   |
| 307, | 272058,   | gp U07173 VCU07173_1, | (HPP 213) | ToxR-activated (tagE) gene ( <i>Vibrio cholerae</i> ) (inner membrane), |   |
| 308, | 23912707, |                       |           |                                                                         |   |
| 309, | 24611590, | sp P30848 PROP_ECOLI, | (HPP 104) | proline/betaine transport protein,                                      |   |
| 310, | 179677,   |                       |           |                                                                         |   |
| 311, | 5083577,  |                       |           |                                                                         |   |
| 312, | 22682813, |                       |           |                                                                         |   |
| 313, | 2843912,  |                       |           |                                                                         |   |
| 314, | 24039587, |                       |           |                                                                         |   |
| 315, | 907827,   |                       |           |                                                                         |   |

FIGURE 1 CONT'D

|      |           |                                                                                                |   |
|------|-----------|------------------------------------------------------------------------------------------------|---|
| 316. | 2035936,  | gp U05676 HRU05676_2, (HPP 12C) vacuolating cytotoxin HpyIori,                                 | 7 |
| 317. | 978477,   | gp L28919 STRFBPSA_1; (HPP 35C) fibrinogen-binding protein FBP54 may be a surface antigen)     |   |
| 318. | 10737627, |                                                                                                |   |
| 319. | 3953143,  |                                                                                                |   |
| 320. | 197166,   | sp P08776 K2C8_XEJLA, (HPP 37H) KERATIN- TYPE II CYTOSKELETAL-Intermediate filament,           |   |
| 321. | 33476715, | sp P16680 PRNA_ECOLI, (HPP 64) alkylphosphonate uptake genes A through Q,                      |   |
| 322. | 14313885, | gp L26016 DVUDCRG_1; (HPP 34G) aspartate chemoreceptor,                                        |   |
| 323. | 34489549, | gi 309688 gp L04161, (HPP 45B) Plasmodium falciparum gametocyte specific antigen,              |   |
| 324. | 3958537,  |                                                                                                |   |
| 325. | 1204418,  | sp P15035 , (HPP 50G) ROD SHAPE-DETERMINING PROTEIN,                                           |   |
| 326. | 5267037,  |                                                                                                |   |
| 327. | 21742157, |                                                                                                |   |
| 328. | 14257751, | sp P15876 MRAY_ECOLI, (HPP 25A) PHOSHO-N-ACETYLMURAMYL-PENTAPEPTIDE-TRANSFERASE,               |   |
| 329. | 25605166, |                                                                                                |   |
| 330. | 23958179, |                                                                                                |   |
| 331. | 4882763,  | (HPP 123) GLUTATHIONE-REGULATED POTASSIUM-EFFLUX SYSTEM PROTEIN,                               |   |
| 332. | 19531291, | sp P03819 , (HPP 123) GLUTATHIONE-REGULATED POTASSIUM-EFFLUX SYSTEM PROTEIN,                   |   |
| 333. | 4455467,  | sp P37388 XYLG_ECOLI, (HPP 454) D-xylose transport atp-binding protein xy19,                   |   |
| 334. | 12969218, |                                                                                                |   |
| 335. | 10312562, |                                                                                                |   |
| 336. | 23475342, |                                                                                                |   |
| 337. | 35397265, |                                                                                                |   |
| 338. | 662933,   | (HPP 453) GLUTATHIONE-REGULATED POTASSIUM-EFFLUX SYSTEM PROTEIN,                               |   |
| 339. | 14864452, | sp P03819 ,                                                                                    |   |
| 340. | 42683,    |                                                                                                |   |
| 341. | 24417212, | (HPP 258) ONE OF THE DIFFERENT ANTIGENIC SEROTYPES OF PROTEIN M,                               |   |
| 342. | 24488537, |                                                                                                |   |
| 343. | 24492192, | sp P08089 ,                                                                                    |   |
| 344. | 41552656, | sp P35640 INVA_BARBA, (HPP 345) invasion protein A,                                            |   |
| 345. | 3157067,  |                                                                                                |   |
| 346. | 12400007, | (HPP 325) glycerolphosphate auxotrophy in plsB background,                                     |   |
| 347. | 21687842, | (HPP 35C) MEMBRANE-ASSOCIATED HYPOTHETICAL 21.7 KD,                                            |   |
| 348. | 34097707, |                                                                                                |   |
| 349. | 30078126, |                                                                                                |   |
| 350. | 33986087, |                                                                                                |   |
| 351. | 5993958,  |                                                                                                |   |
| 352. | 24395801, |                                                                                                |   |
| 353. | 1364378,  | sp P18399 FIXS_RHIME, (HPP 27) homology to NITROGEN FIXATION; TRANSMEMBRANE (Rhizo. meliloti), |   |
| 354. | 11876471, | gp U06471 DN06471_5, (HPP 523) lipopolysaccharide epitope,                                     |   |
| 355. | 625277,   | (HPP 73) variable antigen from Treponema,                                                      |   |
| 356. | 16131887, | sp P16665 ,                                                                                    |   |
| 357. | 14640637, |                                                                                                |   |
| 358. | 22704567, | (HPP 317) FLAGELLAR BASAL-BODY ROD PROTEIN,                                                    |   |
| 359. | 29298130, | (HPP 255) N-ACETYLMURAMATE--ALANINE LIGASE,                                                    |   |
| 360. | 24441412, | sp P17952 ,                                                                                    |   |
| 361. | 14642202, | sp P35652 HPRN_BURSO, (HPP 74) HYPERSENSITIVITY RESPONSE SECRETION PROTEIN,                    |   |
| 362. | 80257,    |                                                                                                |   |
| 363. | 30703183, | sp P15929 FLGH_SALTY, (HPP 247) flagellar basal body L-rim; protein,                           |   |
| 364. | 26054702, |                                                                                                |   |
| 365. | 4787562,  |                                                                                                |   |
| 366. | 23598962, |                                                                                                |   |
| 367. | 234391,   |                                                                                                |   |
| 368. | 34179577, |                                                                                                |   |

FIGURE 1 CONT'D

8

|      |           |                                 |                                                |
|------|-----------|---------------------------------|------------------------------------------------|
| 369, | 34253912, | sp P31134 FLUQ_ECOLI, (HPP 276) | flagellar biosynthetic protein flh2,           |
| 370, | 24882763, |                                 |                                                |
| 371, | 50062,    |                                 |                                                |
| 372, | 1218751,  |                                 |                                                |
| 373, | 4687507,  |                                 |                                                |
| 374, | 23494043, |                                 |                                                |
| 375, | 23442642, |                                 |                                                |
| 376, | 38080063, |                                 |                                                |
| 377, | 4572168,  |                                 |                                                |
| 378, | 6281956,  |                                 |                                                |
| 379, | 23437741, | sp Q03475 LAFB_VIBPA, (HPP 274) | flagellar distal capping protein homolog,      |
| 380, | 25478375, | gp 009868 ECU09868_9, (HPP 10)  | INVOLVED IN F PILUS ASSEMBLY,                  |
| 381, | 10580417, |                                 |                                                |
| 382, | 4035262,  |                                 |                                                |
| 383, | 34099062, |                                 |                                                |
| 384, | 85786,    |                                 |                                                |
| 385, | 17086587, | sp P08150 PBP2_ECOLI, (HPP 105) | D-alanyl-D-alanine carboxypeptidase,           |
| 386, | 22542803, |                                 |                                                |
| 387, | 10723412, | sp P33916 VEUF_ECOLI, (HPP 15)  | hypothetical abc transporter in bcr 5' region, |
| 388, | 11719687, |                                 |                                                |
| 389, | 32234662, |                                 |                                                |
| 390, | 14574201, |                                 |                                                |
| 391, | 40409281, |                                 |                                                |
| 392, | 16281449, | sp P37105 SRP4_BACSU, (HPP 93)  | signal recognition particle protein,           |
| 393, | 36523442, |                                 |                                                |
| 394, | 882827,   |                                 |                                                |
| 395, | 2774062,  |                                 |                                                |
| 396, | 36111066, |                                 |                                                |
| 397, | 31262,    |                                 |                                                |
| 398, | 630,      | sp P15921 , (HPP 426)           | 190KD surface antigen,                         |
| 399, | 392900,   |                                 |                                                |
| 400, | 23646885, |                                 |                                                |
| 401, | 13178562, | gp X72832 SEDEXB_5, (HPP 446)   | stringent response-like protein,               |
| 402, | 4895327,  |                                 |                                                |
| 403, | 21503772, | sp P35162 YX16_BACSU, (HPP 124) | hypothetical protein X,                        |
| 404, | 2001403,  |                                 |                                                |
| 405, | 5878208,  |                                 |                                                |
| 406, | 22303918, | sp P02918 , (HPP 474)           | penicillin binding protein,                    |
| 407, | 4698838,  |                                 |                                                |
| 408, | 2445812,  |                                 |                                                |
| 409, | 22170182, | (HPP 160)                       | H. influenzae lic-1 operon licA-licD genes,    |
| 410, | 16406581, | sp P27841 COPA_ECOLI, (HPP 96)  | MAGNESIUM AND COBALT TRANSPORT PROTEIN,        |
| 411, | 20023400, |                                 |                                                |
| 412, | 30603402, |                                 |                                                |
| 413, | 4095342,  |                                 |                                                |
| 414, | 23728388, |                                 |                                                |
| 415, | 24406401, | sp P31652 NTS1_RAT, (HPP 207)   | serotonin transport protein,                   |
| 416, | 3385833,  |                                 |                                                |
| 417, | 14344378, |                                 |                                                |
| 418, | 32600912, |                                 |                                                |
| 419, | 30283516, |                                 |                                                |
| 420, | 25625192, |                                 |                                                |
| 421, | 30081291, |                                 |                                                |

FIGURE 1 CONT'D

|      |           |                         |           |                                                    |
|------|-----------|-------------------------|-----------|----------------------------------------------------|
| 422. | 5078593.  | sp P23878 FERC_ECOLI.   | (HPP 263) | ferric enterobactin transport protein fepC.        |
| 423. | 23446896. | gi 471729 gpiU056761.   | (HPP 553) | weak similarity to vacA (duplication?).            |
| 424. | 19537968. |                         |           |                                                    |
| 425. | 32663212. |                         |           |                                                    |
| 426. | 34189716. |                         |           |                                                    |
| 427. | 24609593. |                         |           |                                                    |
| 428. | 9954743.  |                         |           |                                                    |
| 429. | 1408.     |                         |           |                                                    |
| 430. | 3930468.  | sp P37169 MVIN_SALTY.   | (HPP 57)  | VIRULENCE FACTOR MVIN.                             |
| 431. | 84691.    |                         |           |                                                    |
| 432. | 13865928. |                         |           |                                                    |
| 433. | 32036462. |                         |           |                                                    |
| 434. | 3953952.  | gi 471729 gpiU056761.   | (HPP 316) | weak vacA similarity.                              |
| 435. | 291700.   |                         |           | penicillin-binding protein 2.                      |
| 436. | 24708129. |                         |           | PERIPLASMIC DIPEPTIDE TRANSPORT PROTEIN PRECURSOR. |
| 437. | 30100332. | sp P23847 DPPA_ECOLI.   | (HPP 332) |                                                    |
| 438. | 4492217.  |                         |           |                                                    |
| 439. | 10745275. |                         |           |                                                    |
| 440. | 31262.    |                         |           |                                                    |
| 441. | 289077.   |                         |           |                                                    |
| 442. | 24818802. | sp Q05605 EXBB_PSEPU.   | (HPP 35)  | biopolymer transport exbB protein.                 |
| 443. | 10353192. | pi1 SIBAG58.            | (HPP 225) | VirB4 homolog.                                     |
| 444. | 24104558. | sp P07893 ATSY_SYNP6.   | (HPP 344) | PROBABLE COPPER-TRANSPORTING ATPASE.               |
| 445. | 3203142.  |                         |           |                                                    |
| 446. | 32144532. |                         |           |                                                    |
| 447. | 4740887.  |                         |           |                                                    |
| 448. | 4548792.  | sp P33231 LCTP_ECOLI.   | (HPP 400) | L-LACTATE PERMEASE.                                |
| 449. | 34658285. |                         |           |                                                    |
| 450. | 4766691.  | gi 459690 gpiL291891.   | (HPP 422) | methyl-accepting chemotaxis protein.               |
| 451. | 36520792. |                         |           |                                                    |
| 452. | 4744128.  | sp P16439 FLAG_SALT.    | (HPP 320) | flagellar basal-body rod proteins                  |
| 453. | 29454837. |                         |           |                                                    |
| 454. | 15039062. |                         |           |                                                    |
| 455. | 4805318.  |                         |           |                                                    |
| 456. | 36594167. |                         |           |                                                    |
| 457. | 785437.   |                         |           |                                                    |
| 458. | 23526667. | sp P07365 CHEW_ECOLI.   | (HPP 84)  | chemotaxis protein chew.                           |
| 459. | 156587.   | gi 459688 gpiL291891.   | (HPP 51)  | transmembrane receptor.                            |
| 460. | 15824052. |                         |           |                                                    |
| 461. | 4578469.  |                         |           |                                                    |
| 462. | 29844512. |                         |           |                                                    |
| 463. | 24415917. | gi L23426 NCOPHOSPHO_1. | (HPP 236) | phosphoglucomutase.                                |
| 464. | 24298127. | sp P31220 YHBC_ECOLI.   | (HPP 244) | PROBABLE ABC TRANSPORTER.                          |
| 465. | 32952.    |                         |           |                                                    |
| 466. | 32422343. |                         |           |                                                    |
| 467. | 23493756. | sp P33231 LCTP_ECOLI.   | (HPP 134) | L-lactate permease.                                |
| 468. | 214812.   |                         |           |                                                    |
| 469. | 1179838.  |                         |           |                                                    |
| 470. | 98131.    | sp Q08382 MOOB_BHICA.   | (HPP 14)  | molybdenum transport system permease.              |
| 471. | 14714687. | gi 495471 gpiU071451.   | (HPP 370) | vacuolating cytotoxin of Hpylori.                  |
| 472. | 3317501.  |                         |           |                                                    |
| 473. | 19541302. |                         |           |                                                    |
| 474. | 23458840. |                         |           |                                                    |

FIGURE 1 CONT'D

475. 2738378, gp|X76422|NSP2A2.1. (HPP 165) penicillin-binding protein 2.  
 476. 22460468, sp|P15933|FLIC\_SALTY. (HPP 303) FLAGELLAR MOTOR SWITCH PROTEIN P.  
 477. 26380318, gp|495471|gp|U07145|. (HPP 325) vacuolating cytotoxin of Hpylori.  
 478. 24803280, gp|L16627|PASPLP123A.2. (HPP 65) outer membrane 10.2K protein.  
 479. 29843937, sp|P10408|SECA\_ECOLI. (HPP 357) PREPROTEIN TRANSLOCASE SECA SUBUNIT.  
 480. 1431462, sp|P35538|FLHB\_BACSU. (HPP 23) FLAGELLAR BIOSYNTHETIC PROTEIN FLHB.  
 481. 34089087, gp|495471|gp|U07145|. (HPP 325) vacuolating cytotoxin of Hpylori.  
 482. 35445843, sp|P37105|SRP4\_BACSU. (HPP 524) signal recognition particle protein.  
 483. 22687687, sp|P10408|SECA\_ECOLI. (HPP 357) PREPROTEIN TRANSLOCASE SECA SUBUNIT.  
 484. 23473437, sp|P35538|FLHB\_BACSU. (HPP 23) FLAGELLAR BIOSYNTHETIC PROTEIN FLHB.  
 485. 23515833, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 486. 30662792, sp|P23445|FLIL\_BACSU. (HPP 12) H<sup>+</sup>-transporting ATP synthase alpha chain homolog.  
 487. 1171928, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 488. 21767890, sp|P05353|VIRB4\_AGR79. (HPP 148) VIRB4 PROTEIN PRECURSOR.  
 489. 4882652, sp|P25536|YHDE\_ECOLI. (HPP 460) E.coli mcd gene Rod shape-determining protein.  
 490. 23539006, sp|P05353|VIRB4\_AGR79. (HPP 148) VIRB4 PROTEIN PRECURSOR.  
 491. 6517192, (HPP 282) flagellar motor switch protein flin.  
 492. 1933, sp|P38489|NFS1\_ECOLI. (HPP 323) OXYGEN-INSENSITIVE NAD(P)H NITROREDUCTASE.  
 493. 22447252, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 494. 14645905, sp|P23445|FLIL\_BACSU. (HPP 12) H<sup>+</sup>-transporting ATP synthase alpha chain homolog.  
 495. 10675632, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 496. 23831562, sp|P25536|YHDE\_ECOLI. (HPP 460) E.coli mcd gene Rod shape-determining protein.  
 497. 32704686, sp|P05353|VIRB4\_AGR79. (HPP 148) VIRB4 PROTEIN PRECURSOR.  
 498. 24816915, (HPP 282) flagellar motor switch protein flin.  
 499. 24219012, sp|P38489|NFS1\_ECOLI. (HPP 323) OXYGEN-INSENSITIVE NAD(P)H NITROREDUCTASE.  
 500. 4897177, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 501. 4486092, sp|P23445|FLIL\_BACSU. (HPP 12) H<sup>+</sup>-transporting ATP synthase alpha chain homolog.  
 502. 21618785, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 503. 16603418, sp|P25536|YHDE\_ECOLI. (HPP 460) E.coli mcd gene Rod shape-determining protein.  
 504. 4551291, sp|P05353|VIRB4\_AGR79. (HPP 148) VIRB4 PROTEIN PRECURSOR.  
 505. 23867687, (HPP 282) flagellar motor switch protein flin.  
 506. 25976418, sp|P38489|NFS1\_ECOLI. (HPP 323) OXYGEN-INSENSITIVE NAD(P)H NITROREDUCTASE.  
 507. 25525277, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 508. 32431687, sp|P25536|YHDE\_ECOLI. (HPP 460) E.coli mcd gene Rod shape-determining protein.  
 509. 4531568, sp|P05353|VIRB4\_AGR79. (HPP 148) VIRB4 PROTEIN PRECURSOR.  
 510. 19720300, (HPP 282) flagellar motor switch protein flin.  
 511. 24413512, sp|P38489|NFS1\_ECOLI. (HPP 323) OXYGEN-INSENSITIVE NAD(P)H NITROREDUCTASE.  
 512. 4570262, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 513. 29500075, sp|P23445|FLIL\_BACSU. (HPP 12) H<sup>+</sup>-transporting ATP synthase alpha chain homolog.  
 514. 3008217, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 515. 134666, sp|P25536|YHDE\_ECOLI. (HPP 460) E.coli mcd gene Rod shape-determining protein.  
 516. 391313, sp|P05353|VIRB4\_AGR79. (HPP 148) VIRB4 PROTEIN PRECURSOR.  
 517. 4726503, (HPP 282) flagellar motor switch protein flin.  
 518. 2617627, sp|P38489|NFS1\_ECOLI. (HPP 323) OXYGEN-INSENSITIVE NAD(P)H NITROREDUCTASE.  
 519. 24495312, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 520. 30082267, sp|P25536|YHDE\_ECOLI. (HPP 460) E.coli mcd gene Rod shape-determining protein.  
 521. 24300682, sp|P05353|VIRB4\_AGR79. (HPP 148) VIRB4 PROTEIN PRECURSOR.  
 522. 25398250, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 523. 23610905, sp|P23445|FLIL\_BACSU. (HPP 12) H<sup>+</sup>-transporting ATP synthase alpha chain homolog.  
 524. 23573294, gp|520402|gp|U03552|. (HPP 714) sensor protein.  
 525. 485375, sp|P25536|YHDE\_ECOLI. (HPP 460) E.coli mcd gene Rod shape-determining protein.  
 526. 1206675, sp|P05353|VIRB4\_AGR79. (HPP 148) VIRB4 PROTEIN PRECURSOR.  
 527. 23567137, gp|520402|gp|U03552|. (HPP 714) sensor protein.

FIGURE 1 CONT'D



11

|      |                                 |                                                           |
|------|---------------------------------|-----------------------------------------------------------|
| 528. | sp P32113 ATVA_ENTFA. (HPP 34C) | POTASSIUM/COPPER-TRANSPORTING ATPASE A.                   |
| 529. | sp P07117 PUTP_ECOLI. (HPP 16C) | SODIUM/PROLINE SYMPORTER.                                 |
| 530. | gi 415692 gp L260151. (HPP 34S) | putative chemoreceptor.                                   |
| 531. | sp P33231 LCTP_ECOLI. (HPP 77)  | L-lactate permease.                                       |
| 532. |                                 |                                                           |
| 533. |                                 |                                                           |
| 534. |                                 |                                                           |
| 535. |                                 |                                                           |
| 536. |                                 |                                                           |
| 537. |                                 |                                                           |
| 538. |                                 |                                                           |
| 539. |                                 |                                                           |
| 540. |                                 |                                                           |
| 541. | sp P22565 LYTB_ECOLI. (HPP 78)  | INVOLVED IN PENICILLIN TOLERANCE-has signal peptide seq.. |
| 542. |                                 |                                                           |
| 543. |                                 |                                                           |
| 544. |                                 |                                                           |
| 545. |                                 |                                                           |
| 546. |                                 |                                                           |
| 547. | sp P26601 . (HPP 9S)            | integral protein in inner membrane,                       |
| 548. | sp Q03523 . (HPP 4K8)           | ACETYLURAMIDYLALANYL-D-GLUTAMATE-DI-MINOPINELATE LIGASE,  |
| 549. |                                 |                                                           |
| 550. | gi 487637 gp U093641. (HPP 35G) | similarity with eukaryotic myosins.                       |
| 551. |                                 |                                                           |
| 552. |                                 |                                                           |
| 553. | sp P10413 HTEG_ECOLI. (HPP 5C1) | heat shock protein C62.5 - chaperonin-ATPase activity,    |
| 554. | gi 155338 gp M829171. (HPP 3S)  | flagellin.                                                |
| 555. | gi 160409 gp M691831. (HPP 508) | surface antigen.                                          |
| 556. | sp Q01465 HREB_BACSU. (HPP 41)  | rod shape-determining protein envB.                       |
| 557. |                                 |                                                           |
| 558. |                                 |                                                           |
| 559. |                                 |                                                           |
| 560. |                                 |                                                           |

FIGURE 1 CONT'D

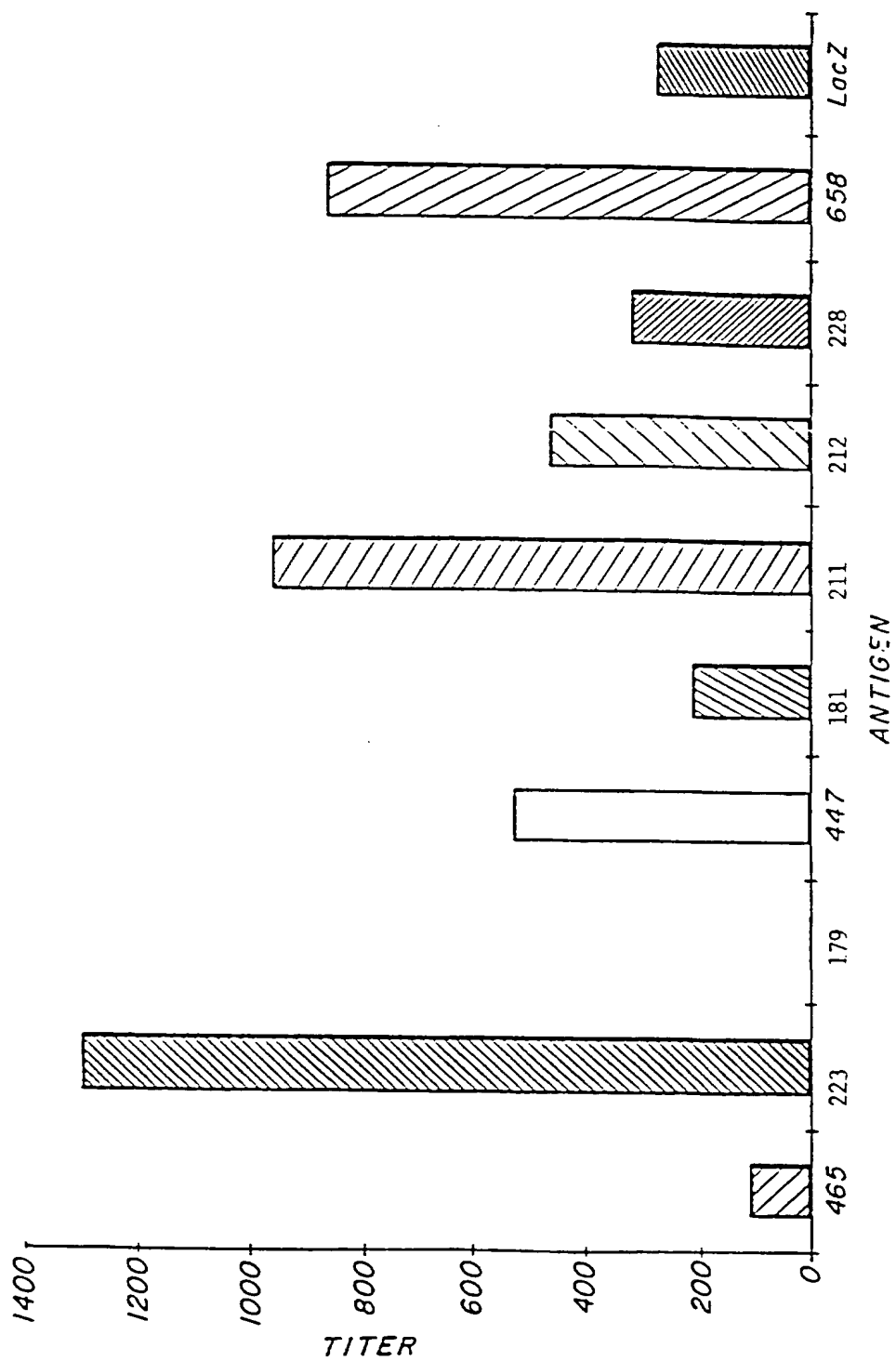


FIGURE 2

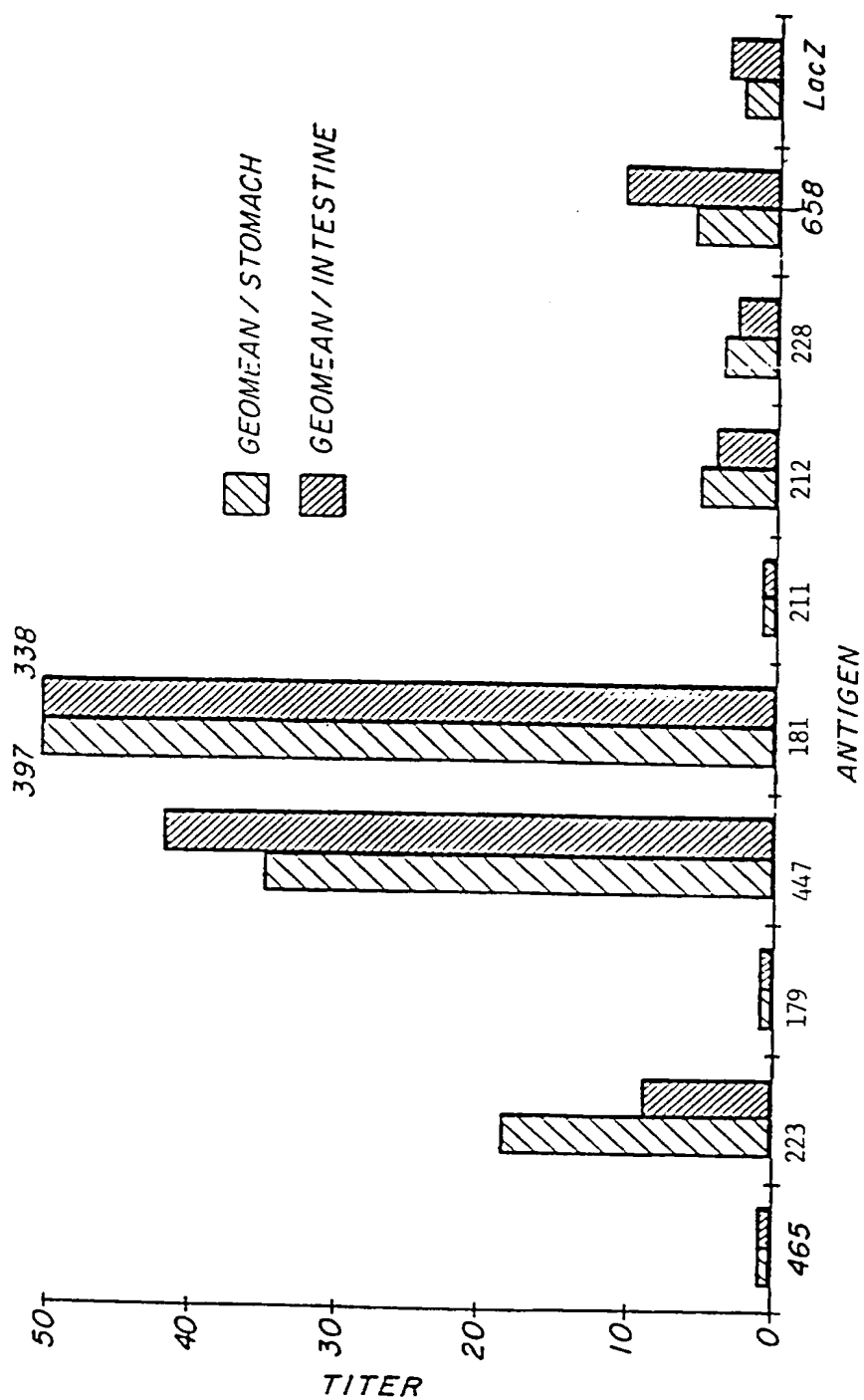


FIGURE 3

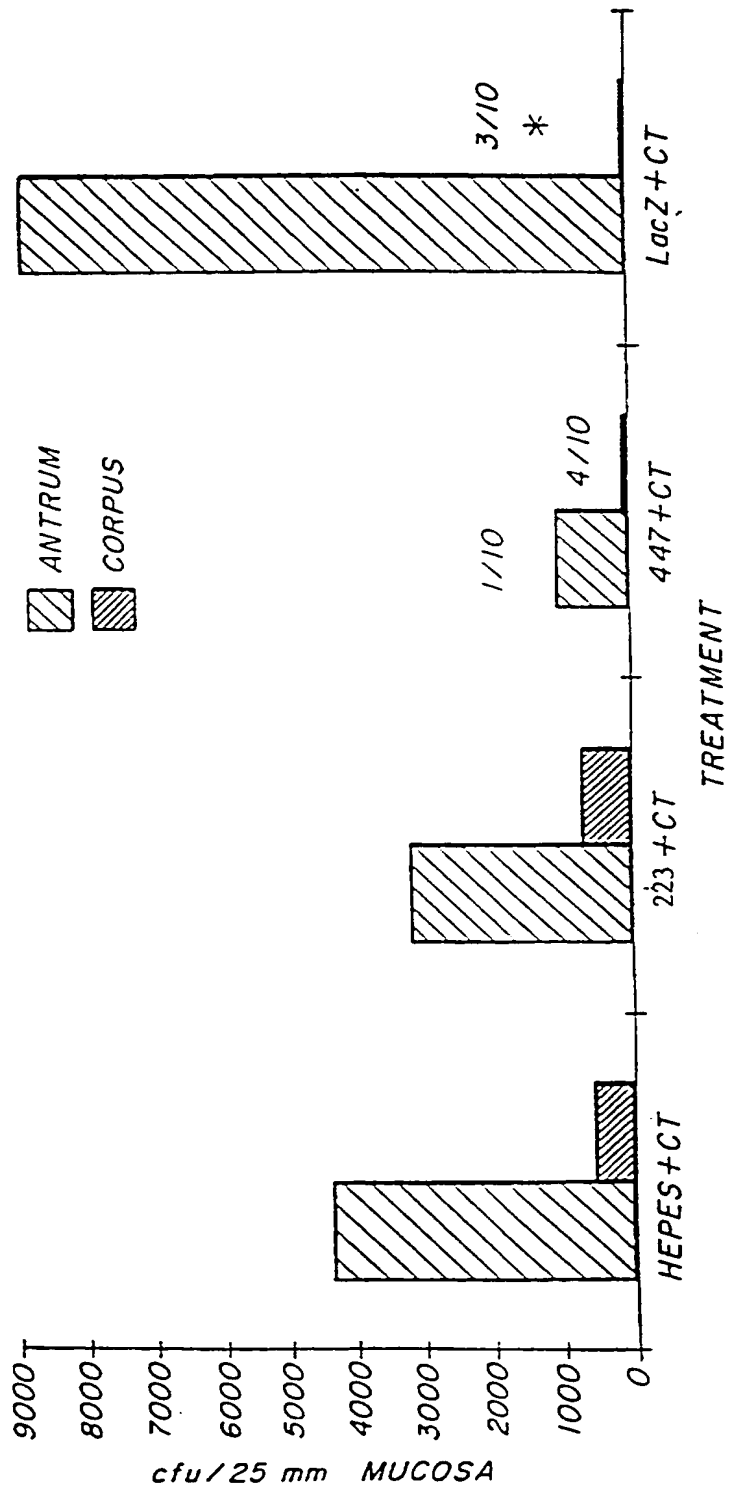


FIGURE 4

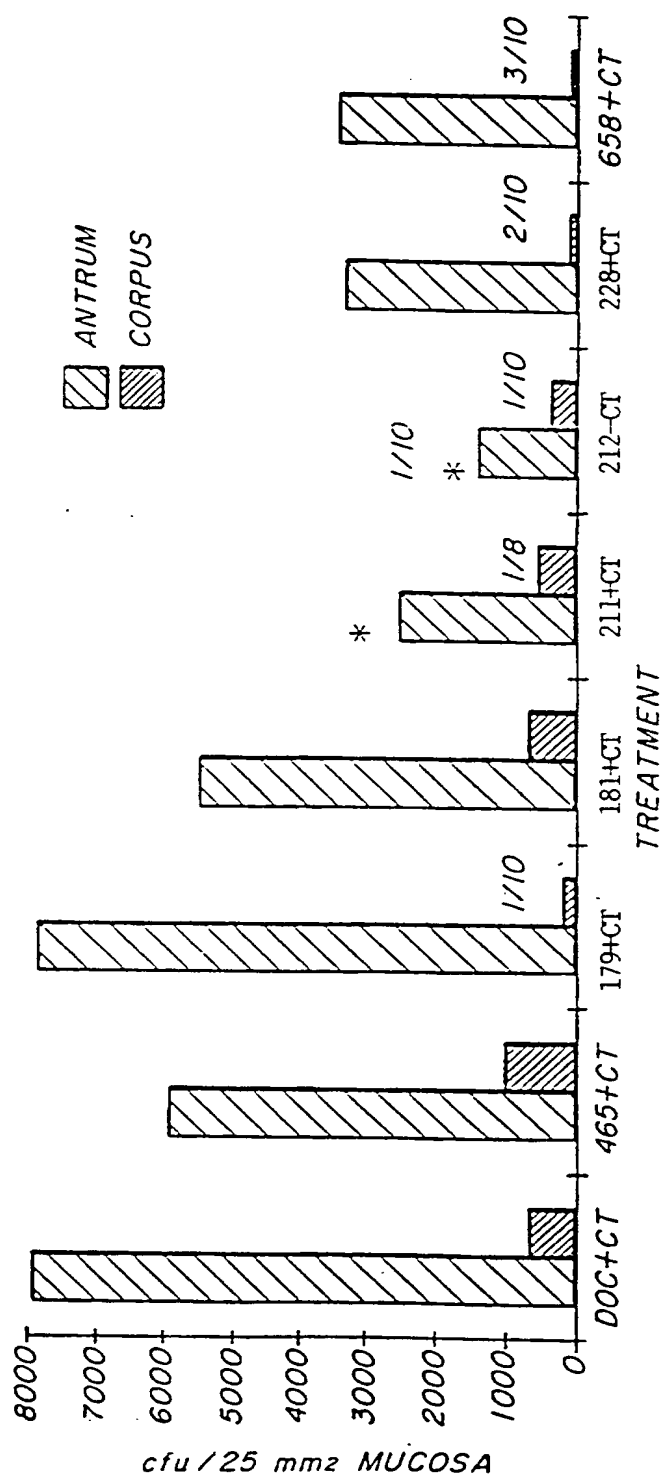


FIGURE 5

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US96/18542

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07H 21/04

US CL :536/23.7

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.7

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, BIOSIS, CA, DERWENT  
search terms: H. pylori, vaccine, gene, protein

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|------------------------------------------------------------------------------------|-----------------------|
| A         | US 5,403,924 A (COVER et al.) 04 April 1995.                                       | 1-7                   |
| A         | US 5,434,253 A (THOMPSON et al.) 18 July 1995.                                     | 1-7                   |
| A, P      | US 5,527,678 A (BLASER et al.) 18 June 1996.                                       | 1-7                   |

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

|                                                                                                                                                                         |                                                                                                                                                                                                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| * Special categories of cited documents:                                                                                                                                | *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention                                              |
| *A* document defining the general state of the art which is not considered to be of particular relevance                                                                | *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone                                                                     |
| *E* earlier document published on or after the international filing date                                                                                                | *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | *G* document member of the same patent family                                                                                                                                                                                                    |
| *O* document referring to an oral disclosure, use, exhibition or other means                                                                                            |                                                                                                                                                                                                                                                  |
| *P* document published prior to the international filing date but later than the priority date claimed                                                                  |                                                                                                                                                                                                                                                  |

Date of the actual completion of the international search

13 MARCH 1997

Date of mailing of the international search report

07 APR 1997

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer  
KENNETH R. HORLICK

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)\*

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US96/18542

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-7 regarding SEQ ID NO: 9, 46, 59, 69, 83, 97, 98, 101, 109, and 114

Remark on Protest

☐  
☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

